What do Bond Investors Learn from Macroeconomic News?

Bruno Feunou Jean-Sébastien Fontaine*
Bank of Canada

Guillaume Roussellet
McGill University

December 2021

Abstract

Macroeconomic data releases drive US bond yields primarily through the term premium instead of the expectation channel. The evidence exploits a monthly specification for yields embedding the impacts of news identified from high-frequency data. To match the facts, we develop and calibrate a no-arbitrage model where investors use data releases with imperfect information to learn about future monetary policy. If macro news carry perfect information, the model predicts that the bonds' Sharpe ratio decreases and the term premium declines by half for every maturity, suggesting that central bank's communication can lower the term premium and financing costs across the economy.

Keywords: Macro-Finance, Term Structure, Imperfect Information, Bayesian Learning
JEL codes: C32, E43, G12

---

* Corresponding author.
– Bank of Canada, jsfontaine@bank-banque-canada.ca, BFeunou@bank-banque-canada.ca.
– Desautels School of Management, guillaume.roussellet@mcgill.ca.
We thank the following individuals for their comments and suggestions: Daniel Andrei, Michael Bauer, Sébastien Bétermier, Anna Cieslak, Greg Duffee, Refet Gurkaynak, Charles Martineau, Jean-Paul Renne, Cisil Sarisoy, Eric Swanson and Jonathan Wright. We also thank participants in the 2019 Southern Economic Association Conference, the 6^{th} CAMF asset pricing workshop, the 2019 Brazilian finance meetings, the 5^{th} Western Annual Financial Econometrics Workshop, the 2019 ESSEC Empirical Finance Workshop, the 9^{th} Bundesbank Term Structure Workshop, and the 2020 World Econometric Summer Conference. This supersedes an earlier draft titled "Long Run Impact of Macro News on Treasury Bond Yields".

1 Introduction

New macroeconomic information is the most important driver of the US bond market. Indeed, the market's responses to the release of new data like inflation, unemployment and output receive considerable attention in economics. The releases have large immediate impacts on the yields, trading volumes, and volatilities of Treasury bonds (Ederington and Lee, 1993; Fleming and Remolona, 1999; Gürkaynak, Sack, and Swanson, 2005). A common intuitive explanation for these large impacts is that bond investors assess the new information about the economy and update what they expect the Federal Reserve to do in the future. In short, macro news act on yields through the expectation channel (e.g., Swanson and Williams 2014). However, it is well known that the yields on long-term bonds embed expectations of future short-term interest rates as well as of returns for bearing risks, the term premium (e.g., Fama 1984; Campbell and Shiller 1991). Since the empirical literature has concluded that this compensation for risk varies with the state of the economy, the intuition that macro news act on yields only through the expectation channel seems incomplete.

This paper is the first to empirically distinguish how new macroeconomic information, as measured in data releases, affects yields through the expectation and term premium channels. To obtain the results, we first rely on high-frequency data to identify the impact of economic releases on bond yields within each month. We then develop a novel specification for bond yields dynamics, set at the monthly frequency, in which we embed the impacts of the data releases, measured in high-frequency data. Mixing these two ingredients delivers new results about the shares explained by macro news in the variances of the expectation and term premium components of yields.

We find that macro news act on bond yields mostly through the responses of the term premiums and less so through the expectation channel. This pattern holds over short horizons following the releases and persists over longer horizons up to several quarters, or years. We provide several robustness checks as well as model-free and consistent evidence based on direct contemporaneous regressions of excess bond returns or changes in survey forecasts. This new stylized fact emerges as a puzzle when looked at through the lens of the intuitive expectation channel. If investors could easily map data releases onto the expected responses of the Federal Reserve, then why is the new information about the economy entering yields largely via the term premium?

We offer an answer that builds on the idea that investors have imperfect information

about the Federal Reserve's response. Intuitively, data releases provide imperfect information about where the economy is heading as well as how the central bank interprets and eventually acts on the new data. One case in point is the 2021 April jobs report, which produced a disappointing increase in the non-farm payrolls. This led to intense speculations on the day of the news around the meaning for future monetary policy. US Treasury Secretary Janet Yellen even held a press briefing to address concerns that looser monetary policy would be inappropriate.[1] By contrast, data releases may contain more precise information about the risks that investors bear from holding bonds and, therefore, about the compensation for these risks embedded in the term premium. For instance, Ludvigson and Ng (2009) document the close links between observable economic variables and future bond returns, consistent with theories implying that investors must be compensated for bearing macroeconomic risks.

To check whether imperfect information quantitatively matches the evidence, we consider a parsimonious no-arbitrage bond pricing model in which investors use the data releases together with Bayes' rule to learn about future central bank policy actions. In the model, news are informative about the expectation component of yields and the compensation for risks, which drives the term premium. We let the information in the data releases have different degrees of precision about these two components of yields.

We calibrate the model parameters to the average level, volatility, correlation, persistence and term premium of bond yields across the maturities. We also calibrate the parameters to the share of the variance of yields attributed to data release. However, we do not use the shares attributed to the term premium or the expectation components in the calibration. The calibrated parameters imply that data releases are much less informative about the central bank's future responses and more informative about the term premium. The model closely matches the variance decomposition statistics that we derived from the data, both at short and long horizons. In addition, the model implies that the macroeconomic information in data releases enters the yields mostly through the term premium.

The calibrated model identifies two sources of risk driving the compensation for risk and the term premium: (i) the variations in the true state of the economy and (ii) the updates in past filtering errors due to the imperfect information. If we eliminate the

[1]See details of Janet Yellen' remark at the link for the Press Briefing. Similarly, CPI growth rates above 2 percent during 2021 triggered a debate between two sides arguing whether or not, or when, the Federal Reserve would accommodate or repress the higher inflation.

filtering errors, by assuming that macro news carry perfect information to investors, the model predicts that the bonds' Sharpe ratios decrease substantially and the term premium is cut by half for every maturity. This exercise suggests that the term premium is a new and distinct channel through which improvements in central bank's communication can lower bond yields and financing costs across the economy.

Our contribution gives researchers the ability to analyze how different sources of economic information enter bond yields and propagate over time, either via the term premium or expectation channel. Existing works focus on the immediate impacts of news, using events studies, and neglect the propagation over time or the distinction between the term premium or expectation channels. Empirically, the results carry important implications for dynamic stochastic equilibrium models that are used to understand the interplay between economic shocks, monetary policy and the business cycles. Mainstream models embed an exclusive role for expectations in transmitting structural shocks to long-term interest rates while only a few cases have that the term premium also channels economic shocks (Rudebusch and Swanson, 2012; Van Binsbergen et al., 2012). Finally, our findings also contribute to the literature looking at whether better communication can make policy more effective by encouraging earlier endogenous bond yields' adjustments to incoming economic data. Blinder et al. (2008) review this "revolution in thinking about central bank's communication" that occurred between an era when central banks were opaque and until the present state with regular and detailed announcements. Our results suggest that improving central bank communications may also reduce the term premium and provide further gains through this new additional channel.[2]

Related Literature

We contribute to a growing literature showing that learning plays a central role in the determination of bond prices. Cieslak (2018) presents evidence that bond market participants do not have full information about the Federal Reserve's monetary policy reaction. Steffensen, Schmeling, and Schrimpf (2021) document persistent investors' errors in expectations of federal responses. Leombroni et al. (2021) find that ECB speeches can

[2]A related channel could be a desire to minimize financial markets' volatility (Stein and Sunderam, 2018). Blinder et al. (2005) review the first FOMC public announcement in February 1994 and note it was introduced by Chairman Greenspan based on the concern that the interest rate hike would disrupt the market. Greesnpan (2001) noted how the growth in stock of outstanding financial instruments made apparent the risk and the deadweight loss associated with opaque policy-making.

have a powerful impact on the yield curve because market participants have imprecise knowledge. We also find evidence of learning by bond investors, but the evidence also shows that new macroeconomic information influences yield mostly through the term premium.

One implication of our results is that investors lean more heavily on information outside of macro data releases to learn and form expectations of what the Federal Reserve will do. For instance, investors may update their beliefs due to information about economic fundamentals that is embedded in equity market valuations (see Cecchetti 2003 for a review). Cieslak, Morse, and Vissing-Jorgensen (2019) suggest systematic informal communication of Federal Reserve officials to the media and financial sector as a channel through which new information about monetary policy has reached the market.

Our results build on a long-standing literature linking high-frequency changes in bond yields to the information in data releases. Gürkaynak, Sack, and Swanson (2005) use the MMS news releases data set to document the responses of short- and long-maturity interest rates to new economic data. We also use the MMS data set and high-frequency data to identify the effect of data releases on yields. Gurkaynak, Kisacikoglu, and Wright (2018) find that information beyond the headline number in the releases also drives a large response.[3] Like them, we use yield changes around the data releases to capture the new information that the report contains, beyond the headline surprise. Using changes in market prices directly as primitives offers the added benefits of controlling for the endogenous variations in investors' attention, which is correlated with the size or volatility of the surprises as well as with the risk premium (Bansal and Shaliastovich, 2011; Andrei and Hasler, 2015; Kacperczyk et al., 2016).

Our empirical approach also accounts for the announcement risk premium earned during small periods of time around the release (Savor and Wilson, 2014; Ai and Bansal, 2018). Focusing on the announcement windows, Wachter and Zhu (2021) explain the announcement risk premium in model where data releases matter for stock prices because they reveal imperfect information to investors concerning underlying shocks that have already occurred. While we also emphasizes that data release contain imperfect information, this growing literature analyzes the unconditional risk-returns relationship around announcement window while we analyze the changes in the conditional expectation and

[3] Beechey and Wright (2009) disentangle the responses in terms of the nominal and real components of interest rates. Faust et al. (2007) and Andersen et al. (2007) analyze the impact of US news on the exchange rate and foreign bond yields.

term premium components. However, like in Wachter and Zhu (2021) our model also embeds the notion that investors learn about the past. This notion is also consistent with low-frequency evidence put forward by Duffee (2021), who finds that asset prices can react to new because investors observe some output growth innovations with a lag.

Altavilla, Giannone, and Modugno (2017) ask how much of the yield variance at the monthly and quarterly frequencies can be attributed to data releases. Like them, we map high-frequency yield responses onto monthly time series. However, we embed these monthly series within a parsimonious dynamic term structure, which means that we can split the yields' responses around the releases between the changes in the expected response of the central bank (the expectation channel) and the changes in expected bond returns above short-term rates (the risk premium channel). Then, we can also analyse the dynamic impacts on both components at horizons between a few months and up to several years.

Our results are distinct from the rich literature analyzing the impact of a specific structural shock on bond yields, an approach pioneered by Rudebusch (1998) and Kuttner (2001) in the case of monetary policy shocks. More recent papers also consider growth shocks and risk premia shocks (e.g., Gurkaynak, Sack, and Swanson 2007; Cieslak and Schrimpf 2018; Bauer and Swanson 2020; Kaminska, Mumtaz, and Sustek 2021). Other recent papers disentangle standard monetary policy shocks and information-revelation shocks (Campbell, Fisher, Justiniano, and Melosi, 2017; Nakamura and Steinsson, 2018; Miranda-Agrippino and Ricco, 2018; Jarociński and Karadi, 2020). In line with our emphasis on the role of learning, Hamilton, Pruitt, and Borger (2011) suggest that FOMC statements provide imperfect information about future policies and Bauer and Swanson (2020) argue that learning by investors is important to understand the response of bond prices to FOMC monetary policy announcements. However, we aggregate together how yields respond to every data release and monetary policy announcement in a month—hence we are commingling the different structural shocks—and we analyze their collective role behind bond yields variations.

There is an extensive literature analyzing asset pricing models with imperfect information.[4] In the context of term structure models, Giacoletti, Laursen, and Singleton (2020)

[4]See Hansen and Sargent (2020) or the review by Pastor and Veronesi (2009). A substantial part of the literature, such as e.g. Andrei, Hasler, and Jeanneret (2019) or Johannes, Lochstoer, and Mou (2016), focuses on learning about the consumption process in structural asset pricing models to help explain the behavior of equity prices.

show that learning about the parameters based on the yield data implies a Bayesian investor with a more variable subjective risk premium and produces accurate yield forecasts. We also find that imperfect information and learning substantially increase the level and volatility of the bond risk premium. However, we analyze learning about the state of economy based on imperfection in the data releases.[5] Hillenbrand (2021) considers a model similar to ours in which bond investors receive imperfect information from either announcements or private signals. However, he focuses on what bond investors learn about the long-horizon path of the interest rate based on the Federal Reserve FOMC announcements. Despite the different focus, he also argues that bond investors face relatively large uncertainty around this path.

The dynamic model that we develop to estimate variance ratios is related to the macro-finance dynamic term structure models (DTSM) studying the linkages between bond prices and macroeconomic variables. Ang and Piazzesi (2003) report the share of yields variance due to innovations in observed macro variables, where the innovations are measured based on a VAR model.[6] Bauer (2015) uses the no-arbitrage restrictions to measure the contemporaneous impacts of different types of macro news jointly across bond yields, but does not explore their role in the long run. We show an explicit mapping onto a generic macro-finance DTSM à la Ang and Piazzesi (2003) with unobserved macroeconomic factors. Intuitively, our approach corresponds to filtering the yields dynamics based on the observed yields factors and the observed yields responses to the data releases. Hence, this approach is related to DTSMs where current yields plus additional information from their history are needed to capture their dynamics.[7]

2 Macro News and the Term Structure

2.1 Bond Yields Variance Ratios

We represent the cross-section of zero-coupon yields with their first three principal components $\mathcal{P}_t$, calculated from yields with yearly residual maturities between one year and ten years, which are available from the Gurkaynak, Sack, and Wright (2006) data set.

[5]Learning about state variables is a slightly different but analytically tractable approach also used in other contexts (e.g., Andrade, Crump, Eusepi, and Moench 2016).

[6]Other leading examples of macro-finance DTSMs include Ang, Piazzesi, and Wei (2006), Diebold, Rudebusch, and Aruoba (2006), Rudebusch and Wu (2008) and Moench (2012).

[7]That is, models in which yields are not "Markovian", see Cochrane and Piazzesi 2005; Joslin, Le, and Singleton 2013; Feunou and Fontaine 2018; Hanson, Lucca, and Wright 2021).

In this representation, the yield with m periods left until maturity $y_t^{(m)}$ is given by:

$$y_t^{(m)} = A_m + B_m^\top \mathcal{P}_t. \tag{1}$$

Equation (1) tells us that the yield curve moves because $\mathcal{P}_t$ moves. The coefficients A_m and B_m can be estimated freely or in a restricted way to embed the no-arbitrage conditions. We choose the latter but this does not influence our results.[8] Note that Equation (1) has no bearing on the dynamics of yields. In particular, the principal components are not necessarily sufficient to forecast yields.

To distinguish the variations in $\mathcal{P}_t$ that are caused by macroeconomic data releases, we introduce the following accounting identity:

$$\Delta\mathcal{P}_t = \Delta\mathcal{P}_{n,t} + \Delta\mathcal{P}_{y,t}, \tag{2}$$

where $\Delta\mathcal{P}_{n,t}$ captures the variations caused by the release of new data and $\Delta\mathcal{P}_{y,t}$ represents the variations occurring at other times and due to other types of information.

Using Equations (1)-(2) and fixing a forecasting horizon $h \geq 1$, we define the share of the variance of the yields $y_{t+h}^{(m)}$ attributed to economic releases:

$$\mathcal{V}_h^{(m)} \equiv \frac{B_m^\top \mathbb{V}\left(\mathcal{P}_{n,t+h}|\underline{\mathcal{F}_t}\right) B_m}{B_m^\top \mathbb{V}\left(\mathcal{P}_{n,t+h} + \mathcal{P}_{y,t+h}|\underline{\mathcal{F}_t}\right) B_m}, \tag{3}$$

where h is the forecast horizon, $\mathbb{V}\left(\mathcal{P}_{n,t+h}|\underline{\mathcal{F}_t}\right)$ is the variance of the news components given the current information set $\underline{\mathcal{F}_t} = \{\mathcal{F}_t, \mathcal{F}_{t-1}, \ldots\}$ that is relevant to forecast yields (defined in the next section). Equation (3) is similar to an R^2 ratio. The conditioning in Equation (3) means that this ratio captures yield variations caused by economic data released between t and $t+h$ and ignores predictable yield variations caused by economic information released before t (e.g., due to a bond risk premium).

To estimate the variance ratios given by Equation (3), we proceed in three steps. We first specify the time-series dynamics of the principal components $\mathcal{P}_t$ to produce forecasts and obtain forecast error variances. Second, we identify the sub-components $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$

[8] Duffee (2018) estimates the coefficients A_m and B_m in Equation (1) with ordinary least squares. The no-arbitrage condition implies restrictions on coefficients across maturities. See, e.g., Joslin, Singleton, and Zhu (2011).

using high-frequency data. Finally, we estimate the model, including the B_m coefficients, and report the variance ratios.

2.2 A Dynamic Macro-Finance Framework

The specification of the yields' dynamics is grounded in a generic macro-finance model à la Ang and Piazzesi (2003):

$$\begin{pmatrix} \mathcal{P}_{t+1} \\ \mathcal{M}_{t+1} \end{pmatrix} = \begin{pmatrix} \Phi_{\mathcal{P}} & \Phi_{\mathcal{P},\mathcal{M}} \\ \Phi_{\mathcal{M},\mathcal{P}} & \Phi_{\mathcal{M}} \end{pmatrix} \begin{pmatrix} \mathcal{P}_{t} \\ \mathcal{M}_{t} \end{pmatrix} + \begin{pmatrix} \omega_{\mathcal{P}} & \omega_{\mathcal{P},\mathcal{M}} \\ 0 & \omega_{\mathcal{M}} \end{pmatrix} \begin{pmatrix} u_{t+1} \\ \Delta\mathcal{P}_{n,t+1} \end{pmatrix}, \tag{4}$$

where $\mathcal{P}_t$ and $\mathcal{M}_t$ have the same dimension (for simplicity); u_{t+1} and $\Delta\mathcal{P}_{n,t+1}$ are uncorrelated; $\mathbb{E}\left(u_{t+1}|\underline{\mathcal{F}_t}\right) = 0$ and $\mathbb{E}\left(\Delta\mathcal{P}_{n,t+1}|\underline{\mathcal{F}_t}\right) = 0$; and where the innovations have variances $\mathbb{V}(u_t) = \Omega_u$ and $\mathbb{V}\left(\Delta\mathcal{P}_{n,t}\right) = \Omega_n$.

The unobserved variable $\mathcal{M}_t$ represents <u>all</u> the information in the history of macroeconomic data that is relevant to forecast bond prices. The question of what relevant variables $\mathcal{M}_t$ should embed remains unsettled in the literature. For instance, Ang and Piazzesi (2003) use one principal component from inflation-related variables and one principal component from real activity variables, but several other choices have been proposed, including the use of lagged values of the macroeconomic variables or the use of a dynamic factor model to summarize a large panel of macroeconomic variables (e.g., Ludvigson and Ng 2009; Joslin, Priebsch, and Singleton 2014; Coroneo, Giannone, and Modugno 2016).

We take a different route to bypass the construction of $\mathcal{M}_t$. One reason for this choice is that we are not interested in using the yield curve to forecast the macroeconomic state variable $\mathcal{M}_t$. Our approach uses high-frequency data to identify the impacts of major data releases on the yield curve. We then aggregate the individual news impact to construct the monthly component $\mathcal{P}_{n,t}$, which we then directly use as the innovation in Equation (4). Then, based on the information set $\mathcal{F}_t = \{\mathcal{P}_t, \Delta\mathcal{P}_{n,t}\}$, we can recover the conditional dynamics of the yield factors $\mathcal{P}_t$,

$$\mathcal{E}_t \equiv \mathbb{E}\left(\Delta\mathcal{P}_{t+1}|\underline{\mathcal{F}_t}\right), \tag{5}$$

by applying the Kalman filter to Equation (4), which leads to:

$$\mathcal{E}_t = K_0 + \Theta\mathcal{E}_{t-1} + K_{\mathcal{P}}\mathcal{P}_t + K_n\Delta\mathcal{P}_{n,t} + K_y\left(\Delta\mathcal{P}_t - \Delta\mathcal{P}_{n,t}\right), \tag{6}$$

where the innovations are given by:

$$\Delta\mathcal{P}_{t+1} - \mathbb{E}\left(\Delta\mathcal{P}_{t+1}|\underline{\mathcal{F}_t}\right) = \Delta\mathcal{P}_{n,t+1} + u_{t+1}. \tag{7}$$

The intuition in Equation (6) is that, by observing the history of the yield principal component $\mathcal{P}_t$ and news releases impacts $\Delta\mathcal{P}_{n,t}$, we can produce the same forecast of future yields as if we could also observe the macro variable $\mathcal{M}_t$.[9] Appendix A.1-A.2 provide the derivation and include the explicit mapping between the parameters of Equations (4) and (6); these two models have the same number of degrees of freedom, as well as conditions for stationarity and invertibility of the model.

2.3 Measuring the Impact of Macro News

To estimate Equation (6), our next step is construct an observable counterpart to $\mathcal{P}_{n,t}$. We do this by aggregating the impact of the major data releases within the month. Our approach uses high-frequency financial market data in the spirit of Fleming and Remolona (1999) and Kuttner (2001). We assume that the movements of yields in a small window around the time of every data release are caused by the new information that investors receive from this release.[10]

We use data about CME futures contracts for delivery of Treasury securities and denote the futures yields by $F^{(\mu)}_{t,\tau}$, where the superscript $\mu \in \{24,\ 60,\ 120\}$ is the maturity of the underlying bond and the index $\tau \in [0, 1)$ measures how much time has passed in the month. We write the high-frequency changes as follow:

$$\Delta F^{(\mu)}_{n,t+\tau_j} := F^{(\mu)}_{t,\tau_j+\mathrm{d}t} - F^{(\mu)}_{t,\tau_j-\mathrm{d}t}\,, \tag{8}$$

where $j \in \{1,\ ...,\ \mathcal{J}_t\}$ indexes the different types of data releases, τ_j is the time of the release and $\mathcal{J}_t$ is the number of releases in month t. Appendix A.3 provides more details about the high-frequency data.

To identify the relevant data releases, we follow the existing literature and use all the scheduled US releases in the MMS data set (see e.g., Kilian and Vega (2011) and Altavilla, Giannone, and Modugno 2017), along with data about the monetary policy

[9] That is, the filtrations generated by $\{\mathcal{P}_t,\ \Delta\mathcal{P}_{n,t}\}$ and $\{\mathcal{P}_t,\ \mathcal{M}_t\}$ are equal.

[10] The timings of the releases are pre-determined and contain no information.

announcements from the FOMC and Piazzesi (2005).[11] This approach captures the most important regular releases but, to be clear, there are other sources of macroeconomic information. For instance, data releases outside of the US, official speeches, and other events influence economic conditions and bond yields. We do not attempt to exhaust all the possibilities. Nonetheless, this standard set of US data releases offers a long time series of exogenous variations in yields, on which investors and the Federal Reserve put great emphasis, and provides a laboratory to test how macro information maps onto bond yields.

The movements in yields attributed to all data releases during month t is given by the sum of all the announcement impacts:

$$\Delta F_{n,t}^{(\mu)} = \sum_{j=1}^{\mathcal{J}_t} \Delta F_{n,t+\tau_j}^{(\mu)} . \tag{9}$$

The aggregation scheme in Equation (9) is a natural choice because news releases that had a larger impact on market prices are given larger weights. By using price changes directly, this scheme captures the wide range of new information that is contained in data releases beyond the headline surprise, as shown by Gurkaynak, Kisacikoglu, and Wright (2018). For robustness, Section 4.4 considers several alternative aggregation schemes.

We then map the monthly variations measured from the futures markets onto the principal components of yields. To motivate this mapping, Appendix A.3 shows that the information spanned by the CME futures yields and the GSW zero-coupon yields is virtually the same.[12] Formally, we construct $\Delta \mathcal{P}_{n,t}^{(i)}$ by running the following regression at the monthly frequency:

$$\Delta \mathcal{P}_t^{(i)} = \underbrace{\beta_i^\top \Delta F_{n,t} - \gamma_i \mathcal{J}_t}_{\Delta \hat{\mathcal{P}}_{n,t}} + \eta_t^{(i)}, \tag{10}$$

where the vector $\Delta F_{n,t}$ stacks the monthly measures across the three maturities.[13]

[11] The number of data releases in our sample gradually increases over time. There were typically around 20 distinct monthly releases after the late 1990s.

[12] Contemporaneous regressions of monthly GSW zero-coupon yields on the three principal components of the futures yields produce R^2s of 0.985 or more. See Table A.2.

[13] Table A.1 in the Appendix reports that R^2s based on Equation (10) range from 18 to 33 percent across the components $\mathcal{P}_{i,t}$.

The term $\gamma_i \mathcal{J}_t$ is a time-varying intercept that cleans our measure of a potential drift in futures prices that may be caused by compensation for the risk of holding the contract around the release time. Savor and Wilson (2014) show that the average excess returns earned around announcement days is significant and aligned with the theoretical predictions from the CAPM. Appendix A.3 provides a formal derivation for the risk premium adjustment. In unreported results, we checked that our results are qualitatively unchanged if we ignore this adjustment. Finally, the appendix shows that the components $\Delta\hat{\mathcal{P}}_{n,t}$ that we recover are white noise, consistent with the assumption in Equation (6), while the raw principal component changes $\Delta\mathcal{P}_t$ are not white noise.

2.4 Model-Implied Variance Ratios

We will present estimates of variance ratios like $\mathcal{V}_h^{(m)}$ in Equation (3) separately for the expectation and term premium components of bond yields:

$$y_t^{(m)} = \underbrace{\mathbb{E}\left[\frac{1}{N}\sum_{i=0}^{m-1} y_{t+i}^{(1)} \,\middle|\, \underline{\mathcal{F}_t}\right]}_{\mathrm{EH}_t^{(m)}} + \mathrm{TP}_t^{(m)}, \tag{11}$$

where $\mathrm{EH}_t^{(m)}$ is the expectation component and $\mathrm{TP}_t^{(m)}$ is the term premium. A common intuition is that the release of new economic data acts on yields through changes in the expectation components $\Delta\mathrm{EH}_t^{(m)}$. Indeed, financial markets experts' commentaries following the news almost exclusively discuss its impact in terms of what the central bank response will be. Changes in nominal bond yields are commonly interpreted as changes in future inflation and policy rates. Mainstream DSGE models also embed this exclusive role that expectations have in transmitting structural shocks to bond prices.[14]

We check this intuition, using the variance ratios to measure the relative importance of the expectation and term premium channels in the transmission of news to bond prices. Because of the linearity assumptions embedded in the model, the variance ratios are available in closed form and can be easily computed once we estimate the parameters of the risk-neutral and historical dynamics. The expectation component $\mathrm{EH}_t^{(m)}$ in Equation (11) has a closed-form expression because the short rate $y_{t+i}^{(1)}$ is a linear combination of the

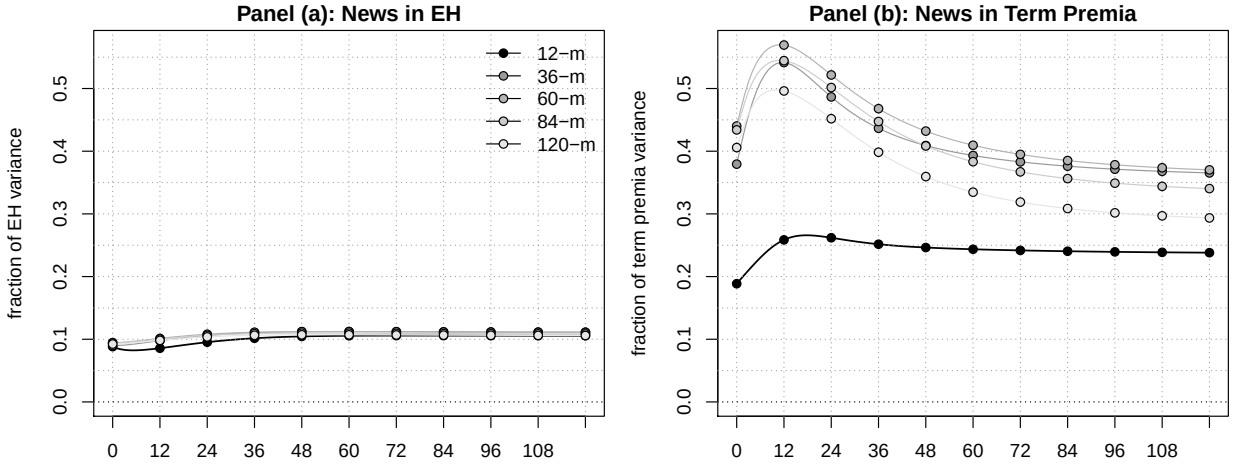
[14] Rudebusch and Swanson (2012) and Van Binsbergen et al. (2012) provide important exceptions where the risk premium also channels macroeconomic shocks to the bond yields.

principal components $\mathcal{P}_{t+i}$ and the conditional expectation is also linear (Equation 6).[15] In the case of the term premium component, because the risk-neutral dynamics belong to the family of affine term structure models, the variance ratio can be computed in closed form as well.

2.5 Results

Estimation of the parameters involves details that we relegate to the appendix. In particular, Appendix A.4 details the estimation of the historical dynamics for $\mathcal{P}_t$ given in Equation (6), provides the likelihood in closed-form and reports the parameter estimates.[16] For the yield coefficients B_ms in Equation (1), we follow Joslin, Singleton, and Zhu (2011) and assume that the bond factors $\mathcal{P}_t$ follow Gaussian VAR(1) dynamics under the risk-neutral measure. We then estimate the risk-neutral parameters by minimizing the squared errors between the observed and fitted yields with 1, 3, 5, 7 and 10 years to maturities.

Figure 1: Share of the Yields Components' Variances Attributed to Data Releases
Panel (a) reports the share of the variance of forecast errors for yields expectation components $\text{EH}_t^{(m)}$ that is attributed to data releases over different horizons. Panel (b) reports this share for the term premium $\text{TP}_t^{(m)}$. Monthly data, 1995-2016.



[15]The expectation component has the following form: $\text{EH}_t^{(m)} = \tilde{A}_m + \tilde{B}_m^\top \mathcal{P}_t + \tilde{C}_m^\top \mathcal{E}_t + \tilde{D}_m^\top \Delta\mathcal{P}_{n,t} + \tilde{E}_m^\top (\Delta\mathcal{P}_t - \Delta\mathcal{P}_{n,t})$.

[16]Our baseline specification (labeled CM-UR) leaves the parameters in Equation (6) unrestricted. Section 4.3 shows that the results are essentially the same in many popular specifications of the yields dynamics that are nested in the baseline specification.

Based on these estimates, Panel (a) of Figure 1 reports the variance ratios for the expectation component EH across bond maturities and forecast horizons while Panel (b) reports results for the term premium. Panel (a) shows that the variance ratio is around 10 percent for the expectation component for every horizon and maturity. Thus, investors seem to update their expectation of future short-term interest rates at other times, outside the windows when the data are released. Instead, Panel (b) shows that data releases are an important driving force of yields via the term premium. The variance ratios attribute around 30 to 40 percent of the term premium variations to data releases across horizons and maturities, except for a hump with higher shares for some of the short horizons.

One plausible reason for the smaller contribution of news to the variance of the expectation component is that investors have imperfect information about the mapping between the releases and the central bank decisions. This can happen in two different ways. One way is that the data releases provide imperfect information about where the economy is heading. The other reason may be that there is imperfect information about the central bank's response to the news. Indeed, Section 3 introduces this type of imperfect information in a no-arbitrage bond pricing model to match the variance ratios. Hence, the variance ratios that we document provide significant evidence and have important implications for equilibrium models of bond prices with imperfect information.

One reason why the variance ratios are relatively high for the term premium component could be that data releases are relatively more informative about the investors' compensation for risk. First, the data releases may be informative about, e.g., consumption risk, long-run risks, habit risk and crash risk. Second, data releases may carry information about the quantity of risk that attracts compensation. For instance, Berger, Dew-Becker, and Giglio (2020) and Dew-Becker, Giglio, and Kelly (2021) show that the realized volatility that is associated with new information carries significant risk premium. Overall, the evidence points to the importance of embedding significant roles for both imperfect information and risk premium channels to understand how new information about the economy determine bond prices.

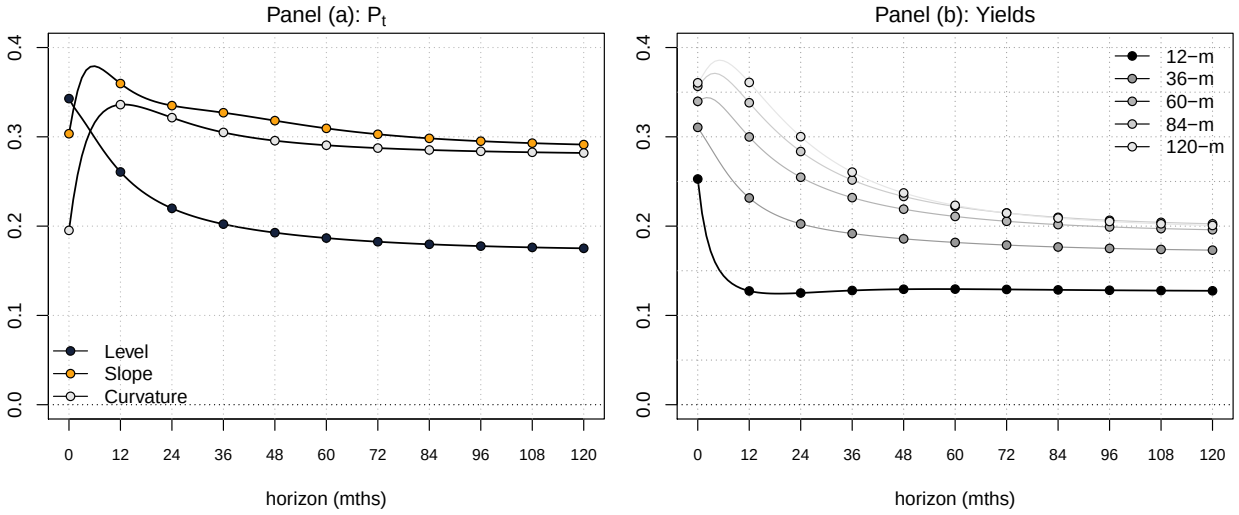
We do not emphasize the absolute levels of the variance ratios. As we noted before, sampling the pre-scheduled data releases could miss many events that contain important macroeconomic information. For instance, speeches, new legislative initiatives, foreign events, etc., also influence bond prices. However, focusing on the pre-scheduled releases identifies the causal impacts of some of the most basic information about the economy.

Observing that this information has a higher impact on the variance of the term premiums is relevant, and possibly puzzling, even if this sample does not capture the population of relevant macro news.

Next, we put together the expectation and term premium components and report results for bond yields in Figure 2. Panel (a) reports the shares for the level, slope and curvature factors separately. Overall, we attribute between 15 and 40 percent of the principal components' variances to the data releases. The variance ratios of the level factor $\mathcal{P}_{1,t}$, which drives the larger share of yields' variations, exhibit a declining pattern across horizons, starting from close to 35 percent at the monthly horizon to less than 20 percent at longer horizons. The share for the slope and curvature factors exhibit hump-shaped patterns but that also declines with the horizon.

Figure 2: Share of the Yields' Variances Attributed to Data Releases

Panel (a) reports the share of the variance of the forecast errors for the principal components $\mathcal{P}_t$ that can be attributed to macroeconomic data releases across different forecast horizons. Panel (b) reports this share for yields. Monthly data, 1995-2016.



The shapes in Panel (a) translate into similar patterns for yields in Panel (b) of Figure 2. The variance ratio ranges between 25 and 35 percent across maturities at the monthly horizon. The results at this short horizon are reassuring and consistent with existing regression-based results in Altavilla, Giannone, and Modugno (2017). At the longest horizons, around 15 to 20 percents of yields' variances is attributed to data releases. This declining pattern arises because the variance ratios in Panel (b) mix the variance ratios for the expectation and the term premium components in Figure 1. The variance ratios are

higher for the term premium component and drive the ratios for yields for short forecast horizons. However, the expectation component is more persistent and its ratios drive the results for yields in the case of long forecast horizons.

Overall, the results are surprising in one key aspect. Despite the common intuition that the data releases influence yields mostly (if not exclusively) because investors update how they think the Fed will respond to the new information, we find that most of the yields' variances around the news are due to changes in the term premium. This suggests that investors cannot easily map data releases onto the expected responses of the Federal Reserve and that this source of imperfect information is an important feature of the bond market. This is the mechanism that we explore in the next section.

3 A Model with Imperfect Information

To explain the patterns in empirical variance ratios, we propose a mechanism where macroeconomic releases provide imperfect information about future interest rates and compensation for risks. For instance, the weekly employment report provides a signal about the state of the labor market, but the implications for future monetary policy and the market prices of risks could be unclear. We develop a theoretical no-arbitrage model with imperfect information. Consistent with this mechanism, we find that periods with high uncertainty around monetary policy are also characterized by a low share of yields variances explained by data releases. We then show in a calibration that this mechanism can explain the patterns in the variance ratios in Figures 1 and 2.

3.1 A Two-factor Model

The economic information relevant for the bond market is summarized by two state variables $\mathcal{P}_{1,t}$ and $\mathcal{P}_{2,t}$ with dynamics given by:

$$\mathcal{P}_{i,t} = \mu_i + \phi_i \mathcal{P}_{i,t-1} + \varepsilon_{i,t}, \tag{12}$$

where $|\phi_i| < 1$ and $\varepsilon_{i,t} \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_i}\right)$ are independent white noise. We assume that bond investors do not directly observe the state variables $\mathcal{P}_{i,t}$ at date t but receive a noisy signal $z_{i,t}$ given by:

$$z_{i,t} = \mathcal{P}_{i,t} + \eta_{i,t}, \tag{13}$$

where the noise $\eta_{i,t} \sim \mathcal{N}\left(0, \sigma^2_{\eta_i}\right)$ are uncorrelated with each other and with $\varepsilon_{i,t}$. Investors update their beliefs rationally using Bayesian updating and we denote their best evaluation of the factors with $\mathcal{P}_{t|t} \equiv \mathbb{E}[\mathcal{P}_t|\underline{\mathcal{F}_t}]$.

We interpret z_t as the public but imperfect information contained in the economic data being released throughout month t. There are several intuitive reasons for this assumption: *(i)* the aggregated data can mask heterogeneity across the economy, *(ii)* several of the data releases are provisional numbers that will eventually be revised, *(iii)* the data releases provide a snapshot of the current situation but offer incomplete information about what investors should forecast and, most importantly, *(iv)* the investors still have to infer how the central bank will interpret the data and respond to the new information (Stein and Sunderam, 2018), given there uncertainty around the impact of policies the economy (Brainard, 1967).

For simplicity, we assume that bond investors obtain complete information about $\mathcal{P}_{i,t}$ after one month has passed, at date $t+1$. Therefore, the investors' information set at date t is given by $\underline{\mathcal{F}_t} = \{z_t, \mathcal{P}_{t-1}, z_{t-1}, \mathcal{P}_{t-2}, \ldots\}$ and Appendix A.5.1 describes the corresponding filtering problem and derives the state-space representation of $\mathcal{P}_{t|t}$.

The assumption that information about economic conditions at date t is revealed after one period is similar to the approximation made by Stein and Sunderam (2018) to solve a dynamic version of their equilibrium model with monetary policy gradualism. Allowing for the for values $\mathcal{P}_t$ to be revealed more slowly would produce a richer information set and the model may have more flexibility to fit the data, but at the costs of parsimony and the loss of clarity and tractability.

3.2 Bond Yields

The pricing kernel in this economy is given by $\log(\mathrm{M}_{t+1}) = -r_{t|t} + \log(\mathrm{M}_{1,t+1}) + \log(\mathrm{M}_{2,t+1})$, where $r_{t|t}$ is the short-term rate given the time-t information set of investors. Each component $\mathrm{M}_{i,t+1}$ for $i = 1, 2$ is given by:

$$\log(\mathrm{M}_{i,t+1}) = \lambda_{i,t}\left[z_{i,t+1} - \mathbb{E}\left(z_{i,t+1}\middle|\underline{\mathcal{F}_t}\right)\right] - \frac{1}{2}\lambda^2_{i,t}\left(\sigma^2_{\varepsilon,i}\left[1 + \phi_i^2\left(1 - \mathcal{K}_i\right)\right] + \sigma^2_{\eta,i}\right), \tag{14}$$

where $\lambda_{i,t}$ is the price of risk and $\mathcal{K}_i = \frac{\sigma^2_{\varepsilon_i}}{\sigma^2_{\varepsilon_i} + \sigma^2_{\eta_i}}$ is the Kalman gain measuring the precision of the information about $\mathcal{P}_{i,t}$ that is revealed by $z_{i,t}$. The innovations $z_{i,t+1} - \mathbb{E}(z_{i,t+1}\,|\,\underline{\mathcal{F}_t})$

in Equation (14) mix two sources of risk: the unexpected content of the releases $z_{i,t+1}$ and the correction of the filtering errors $\mathcal{P}_{i,t} - \mathcal{P}_{i,t|t}$. We attribute the same price of risk to both sources of information, for simplicity and because both concern the same state variables $\mathcal{P}_{t+1}$. Finally, we have $\mathbb{E}_t[\mathrm{M}_{t+1}|\underline{\mathcal{F}_t}] = e^{-r_{t|t}}$, as required.

To ease the interpretation, we give distinct roles to $\mathcal{P}_{1,t|t}$ and $\mathcal{P}_{2,t|t}$ in constructing bond prices. First, we assume that the short-term interest rate is determined by $\mathcal{P}_{1,t|t}$ alone:

$$r_{t|t} = \mathcal{P}_{1,t|t}. \tag{15}$$

Therefore, $\mathcal{P}_{1,t|t}$ summarizes both the current level and the dynamics of the short rate.[17] Next, $\mathcal{P}_{2,t|t}$ captures investors' time-varying preference toward risks that influence the term premium in bond yields. For this purpose, and following Duffee (2002), the prices of risk are linear in the state variables:

$$\begin{pmatrix} \lambda_{1,t} \\ \lambda_{2,t} \end{pmatrix} = \begin{pmatrix} \lambda_{0,1} \\ \lambda_{0,2} \end{pmatrix} + \begin{pmatrix} \Lambda_1 & \Lambda_{1,2} \\ 0 & \Lambda_2 \end{pmatrix} \begin{pmatrix} \mathcal{P}_{1,t|t} \\ \mathcal{P}_{2,t|t} \end{pmatrix}. \tag{16}$$

The exclusion restriction is included for parsimony and simplicity. Given the assumptions introduced so far, Appendix A.5.2 shows that bond yields $y_t^{(m)}$ satisfy the following no-arbitrage relation:

$$y_t^{(m)} = A_m + B_{1,m}\,\mathcal{P}_{1,t|t} + B_{2,m}\,\mathcal{P}_{2,t|t}\,, \tag{17}$$

where A_m, $B_{1,m}$ and $B_{2,m}$ are given by closed-form recursions.

3.3 Variance Ratios

The variance of $\mathcal{P}_{i,t+h|t+h}$ for any horizon h is given by:

$$\mathbb{V}_{i,h} := \mathbb{V}\left(\mathcal{P}_{i,t+h|t+h}\,|\,\underline{\mathcal{F}_t}\right) = \underbrace{\mathcal{K}_i^2\Sigma_i\left(\frac{1-\phi_i^{2h}}{1-\phi_i^2}\right)}_{\text{releases}} + \underbrace{\mathcal{K}_i(1-\mathcal{K}_i)\Sigma_i\left(\phi_i^2\frac{1-\phi_i^{2h}}{1-\phi_i^2}\right)}_{\text{error correction}}, \tag{18}$$

[17] In a more general multivariate setting, one linear combination of a state vector would represent the current rate and a distinct combination would represent its expected path (e.g., $\delta^\top\mathcal{P}_{t|t}$ and $\delta^\top\phi\mathcal{P}_{t|t}$, respectively in the common notation). Although more realistic, it does not produce additional insight for our purpose.

where $\Sigma_i = \sigma^2_{\varepsilon_i} + \sigma^2_{\eta_i}$ (see Appendix A.5.3). These two terms represent distinct sources of variability in $\mathcal{P}_{i,t|t}$ from the investors' point of view. The first term represents the variability due to data releases. The second term represents the correction of the filtering errors. This term decreases when the signal is more precise, and it disappears if the data releases provide complete information about the state $\mathcal{P}_{i,t}$ (i.e., if $\mathcal{K} = 1$). We then get to an important intermediate result. The share of $\mathcal{P}_{i,t|t}$ variance attributed to economic data releases is given by:

$$\mathcal{V}_i := \frac{\mathcal{K}_i}{\mathcal{K}_i + (1 - \mathcal{K}_i)\phi_i^2}, \tag{19}$$

which varies between 0 and 1 with the precision of the signal. This ratio is constant across the forecasting horizons.

Therefore, the model predicts that the precision of the information contained in the data releases determines the share of the variance explained by these news releases. As a check of this mechanism, we compare the share of yields' realized volatility attributed to data releases each month with the monthly Baker-Bloom-Davis (BBD) monetary policy uncertainty index (MPU) for the United States (Baker, Bloom, and Davis, 2016). In a first step, we compute the sum of squared changes in the 2-year yield observed around data releases within each month, which we normalize with the sum of all squared daily yield changes during that month. We report the 12-month moving average of this realized variance ratio together with the 12-month moving average of the MPU index in Figure 3.[18]

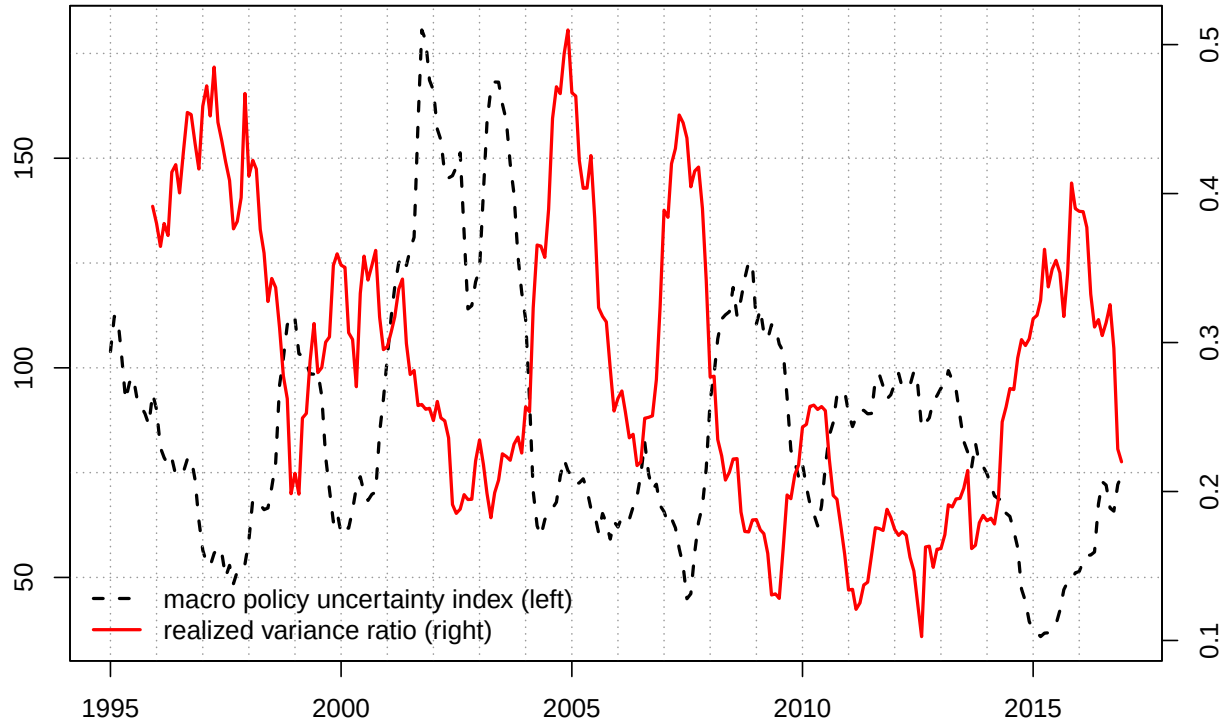
The results show a clear negative relationship between the MPU index and our realized variance ratio. The contribution of the data releases to the realized volatility of yields is high when the MPU index is low (the sample correlation is -0.52). Since the MPU index measures how frequently uncertainty is mentioned in newspapers' discussion of monetary policy, a low value of the index also likely means that the main data releases provide precise information about the economy. Therefore, the relationship shown in Figure 3 is consistent with Equation (19), which predicts that the variance ratio varies with the precision of the news.

Overall, this simple reduced-form exercise provides an initial validation of the mechanism at the heart of our model. While Figure 3 also suggests that economic shocks driving

[18]We focus on the 2-year yield but the results are very similar if we use the 5- or 10-year yield. The realized variance ratios are highly correlated: the first principal component explains 93 percent of the total variations. The results are also similar whether we use the BBD MPPU index based on the broad Access World News set of US newspapers or based on the 10 major newspapers.

Figure 3: Realized Variance Ratio and Monetary Policy Uncertainty

Time-series of (i) the Baker-Bloom-Davis Monetary Policy Uncertainty index (MPU) derived from the Access World News database and available from the Economic Policy Uncertainty web site; and (ii) the share of the realized volatility in the 2-year yield within the month attributed to the data releases. For both variables, we report a 12-month moving average. Monthly data 1995-2016.



the bond yields have conditional variances that vary over time, we leave this for future research. If the variance changes over time, the procedure in Section 2.4 belongs to the family of quasi-maximum likelihood estimators and produces consistent estimates of the parameters as well as of the unconditional variance ratios in Figure 1-2. In the following, we check if the model can quantitatively match these unconditional moments across bond maturities and across horizons.

3.4 The Term Structure of Variance Ratios

The share of bond yields variance explained by macroeconomic releases is a weighted average of the primitive ratios $\mathcal{V}_i$ associated with the $\mathcal{P}_{1,t}$ and $\mathcal{P}_{2,t}$ as follows:

$$\mathcal{V}_h^{(m)} = \omega_h^{(m)} \mathcal{V}_1 + \left(1 - \omega_h^{(m)}\right) \mathcal{V}_2 \,, \tag{20}$$

where the weights $\omega_h^{(m)}$ depend on the horizon h and the maturity m:

$$\omega_h^{(m)} := \left(1 + \frac{B_{2,m}^2}{B_{1,m}^2} \times \frac{\mathbb{V}_{1,h}}{\mathbb{V}_{2,h}}\right)^{-1} . \tag{21}$$

Thus, the ratio of both factors' variances $\mathbb{V}_{1,h}/\mathbb{V}_{2,h}$ determines the pattern that these weights $\omega_h^{(m)}$ exhibit across the horizons h and the ratio of factors loadings $B_{2,m}^2/B_{1,m}^2$ determines the pattern across the maturities m. These patterns are important because the set of weights $\omega_h^{(m)}$ determines the shape of the variance ratios for yields in Figure 2. Since the variance ratios are relatively lower for the expectation component and higher for the term premium component, the model can match the evidence only if these weights decline with the horizon h. Proposition 1 summarizes how the model parameters determine these patterns.

Proposition 1

(i) The weight $\omega_h^{(m)}$ decreases with the bond maturity m.

(ii) The weight $\omega_h^{(m)}$ increases with the forecast horizon h only if $\phi_1 > \phi_2$, and vice versa.

Appendix A.5.4 provides the proof. The first result in Proposition 1 says that the role of the expectation component in the variance of yields decreases with the maturity of the bond. This happens for two reasons. First, the variability of the expectation component decreases with the maturity of the bond because the short rate is stationary and forecasts of the short rate converge to a constant. Second, the term premium varies more at longer maturities because interest-rate risk is higher.

The second result in Proposition 1 says that the importance of the more persistent component of bond yields increases with the horizons. Indeed, the variance of forecast errors increases more rapidly with the horizons for a persistent variable. From the lens of

our model, the case where $\phi_1 > \phi_2$ means that the expectation component of bond yields is more persistent than the term premium, which is an assumption that is supported by the data. This is also the relevant case to match the variance ratios observed in the data. Using Equation (20), the variance ratio decreases at longer horizons if:

$$\mathcal{V}_{h+1}^{(m)} < \mathcal{V}_h^{(m)} \Longleftrightarrow \left(\omega_{h+1}^{(m)} - \omega_h^{(m)}\right)(\mathcal{V}_1 - \mathcal{V}_2) < 0. \tag{22}$$

Figure 1 shows that the variance ratio $\mathcal{V}_1$ of the expectation component is lower than the ratio $\mathcal{V}_2$ for the term premium. Therefore, the variance ratios $\mathcal{V}_h^{(m)}$ decreases with the horizons only if the weight $\omega_h^{(m)}$ increases with the horizon, and this occurs only if the expectation factor is more persistent $\phi_1 > \phi_2$ (see Proposition 1). We conclude that our theoretical model can reproduce the pattern in our empirical findings based on mild and realistic assumptions. We analyse the quantitative implications of the model below.

3.5 Calibration Results

To check the predictions of the model for the variance ratios, we calibrate the parameter to important statistical moments of bond yields. Since the variance ratios are unconditional moments, we calibrate the model to the following set of unconditional moments for the 1- and 10-year yields. Specifically, we match the average, auto-correlation and volatility of each yield. We also match the correlation between these yields, the average term premium and the variance ratios at the one-month horizon $\mathcal{V}_1^{(m)}$. Note that the one-month ratios do not rely on the dynamic models that we estimated earlier. These ratio can be directly obtained from the regression in Equation (10). In addition, these moments do not include any direct information about how the expectation and the term premium components of yields respond to data releases.

Figure 4 reports these statistics for the targeted maturities and shows that the model matches the same moments for several other maturities (3-month, 3-year, 5-year and 7-year maturities). The red dots indicate the moments used for the calibration, which the model matches nearly exactly by design. The yellow boxes indicate the same moments but for other maturities and shows that this simple 2-factor model provides a good fit.[19]

[19]It would also be unrealistic to use a two-factor model to capture conditional moments of yields. Indeed, Section 2.2 shows that capturing the rich dynamic properties of yields *and* the patterns in the variance ratios requires a more general approach than the common VAR(1) dynamics, even with three factors.

Figure 4: Calibration of the Two-Factor Model with Imperfect Information

Calibration of the two-factor model in Section 3.1 to moments of bond yields. The model parameters are calibrated to exactly match the moments indicated by the red dots. Panel (a): average yield curve; Panel (b): volatility curve; Panel (c): cross-correlations; Panel (d): auto-correlations; Panel (e): 1-month variance ratios $\mathcal{V}_1^{(m)}$; Panel (f): term premium.

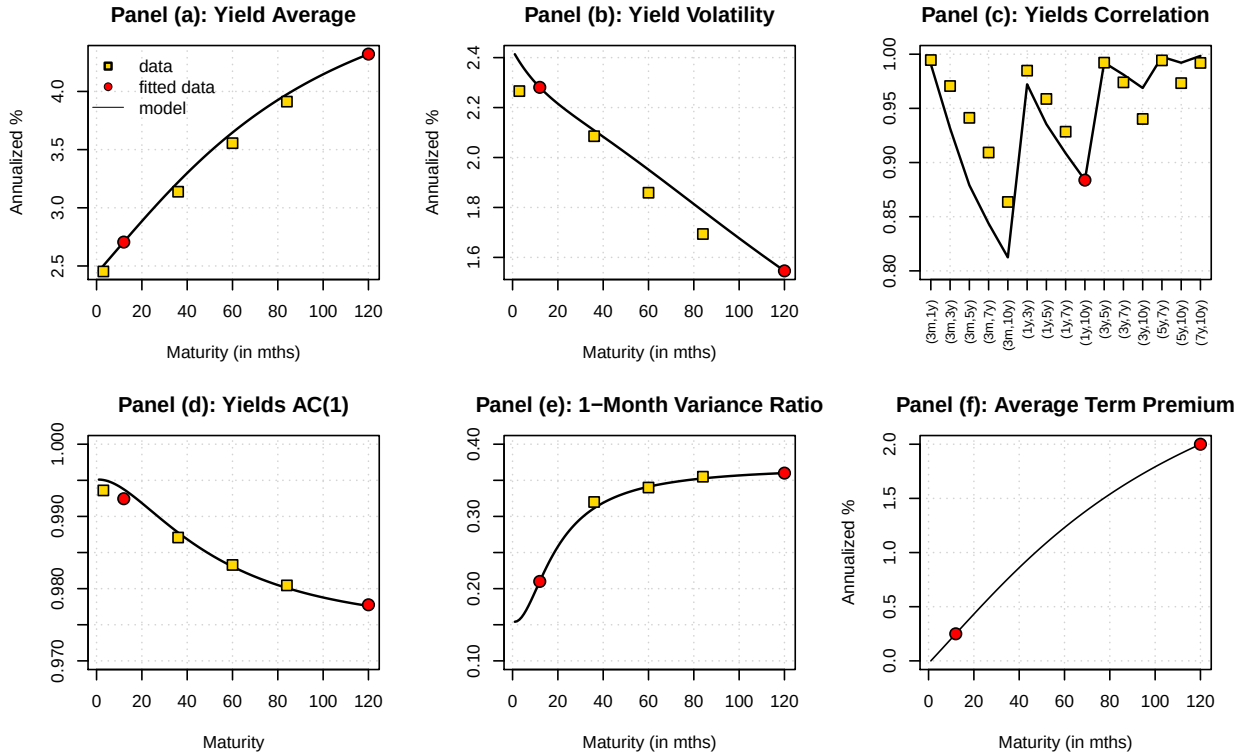


Table 1 reports the calibrated parameters values.[20] While the calibration exercise pins down all parameters simultaneously, different moments are naturally more informative about different subsets of parameters. The persistence and volatility parameters are largely driven by the moments in Panels (a)-(d). We find that the expectation component is more persistent and has more volatile innovations than the term premium. Second, the noise parameters are largely driven by the fact that the 1-month variance ratio increases with the maturity in Panel (e) noise. The parameters imply that the data releases are more informative about the variations in the term premium than that in the expectation component. The standard deviation σ_{η_1} of noise in the expectation component is roughly ten times larger than the standard deviation σ_{η_2} for the term premium. Finally, the parameters of the

[20]Note that we report the parameters of the factor dynamics under the risk-neutral measure, but Appendix A.5.2 reports the one-to-one mapping with the parameters of the prices of risk that we defined in Equation (16).

risk-neutral dynamics are largely determined by the average term premium in Panel (f).

Table 1: Calibrated Parameter Values

Parameter values for two-factor model with imperfect information described in Section 3.1. The parameters were calibrated to unconditional moments of bond yields shown in Figure 4.

		μ	Φ		σ_ε	σ_η	$\mathcal{K}$
$\mathbb{P}$	$\mathcal{P}_{1,t\|t}$	0.0120	0.9951	0	0.2396	0.5640	0.1529
	$\mathcal{P}_{2,t\|t}$	-0.0053	0	0.9482	0.0433	0.0567	0.3684
$\mathbb{Q}$	$\mathcal{P}_{1,t\|t}$	0.0120	0.9870	-0.6321			
	$\mathcal{P}_{2,t\|t}$	-0.0053	0	0.9608			

Equipped with values for all the model parameters, we report the variance ratios implied by the model. Based on the discussion of Proposition 1, the combination of a relatively noisier and more persistent expectation component should drive a decreasing pattern across the horizons in the variance ratios $\mathcal{V}_h^{(m)}$. Figure 5 reports the results. Panel (a) provides the model-implied variance ratios for yields and corresponds to the results in Figure 2. The comparison immediately shows that the calibration reproduces the decreasing pattern that we documented in the variance ratios, although no information about the decline in the variance ratios has been used in the calibration. Hence, these standard statistics of yields are enough to generate the desired patterns once we allow for the learning mechanism.

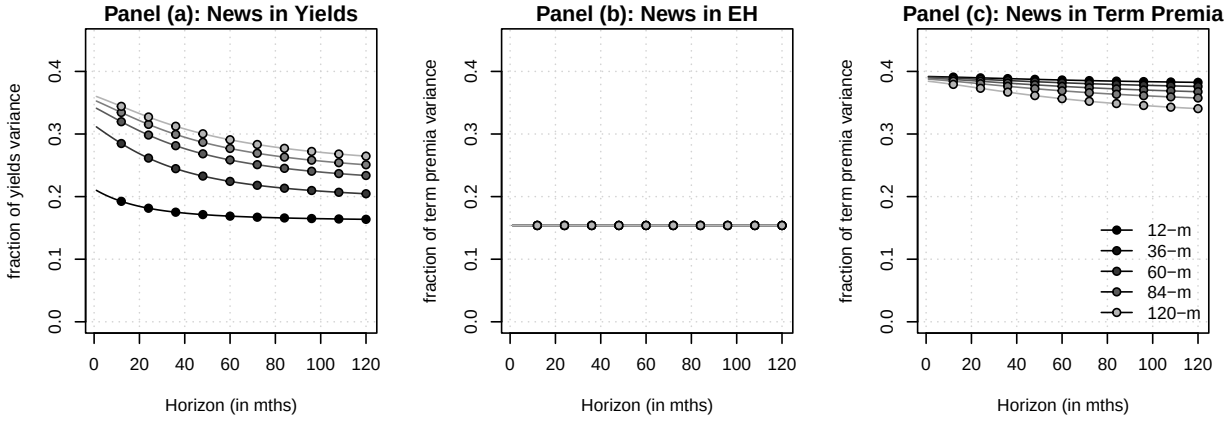
Panels (b)-(c) of Figure 5 reports the model-implied variance ratios for the expectation and term premium components, respectively. These ratios can be compared with the results in Figure 1. In the calibration results, the variance ratios of the expectation component are close to 10 percent and constant, while the variance ratios of the term premium component are close to 40 percent and exhibit a declining pattern. Again, no information about the variance ratios of the expectation or term premium components were used to calibrate the model. The outcome followed from matching standard moments of yields to a model that incorporates learning.

3.6 Model Counterfactuals

The calibration shows that imperfect information can explain the patterns that we observe in variance ratios. In this section, we use the model to assess the importance

Figure 5: Two-Factor Model with Imperfect Information—Variance Ratios

Variance ratios implied from the two-factor model with imperfect information described in Section 3.1 and calibrated to moments of bond yields shown in Figure 4. The variance ratios measure the share of variability attributed to the new information in data releases. Panel (a): ratios for bond yields; Panel (b): ratios for the expectation component of yields; Panel (c): ratios for the term premium component.



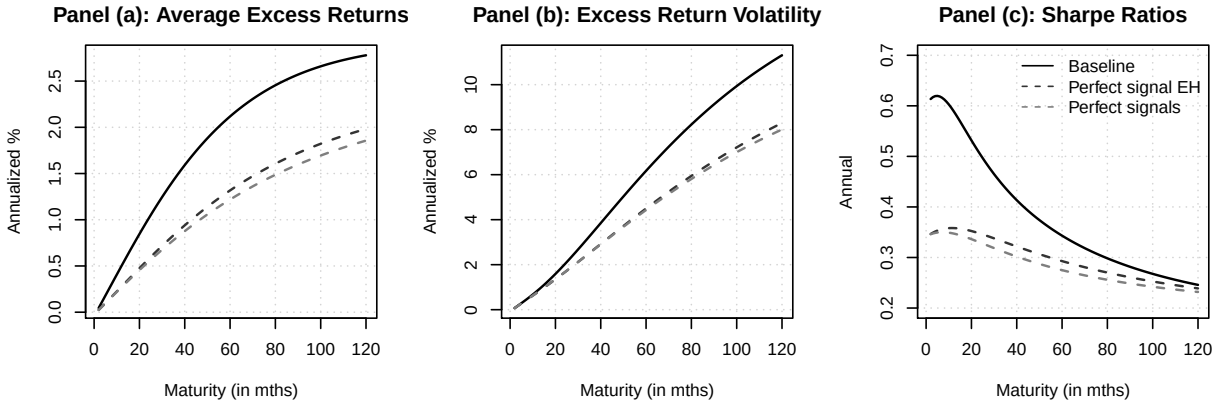
of imperfect information in the risks that bond investors face and the expected returns that bonds offer. We consider the counterfactual case where data releases provide perfect information about the expected path of future short rates (i.e., $\sigma_{\eta_1} = 0$). For completeness, we also consider a perfect information economy, where the data releases convey perfect information about both factors (i.e., $\sigma_{\eta_1} = \sigma_{\eta_2} = 0$). In essence, we keep unchanged the variability in the true state of the economy, but we remove imperfect information completely. This lowers the term premium by reducing or eliminating the risk due to the correcting of past filtering errors.

Panel (a) of Figure 6 reports the average annualized expected bond excess returns—the risk premium—in the baseline and counterfactual calibrations. In the baseline, the risk premium ranges from zero at the shortest maturity to around 2 percent for the 5-year bond and to slightly more than 2.5 percent for the 10-year bond. By contrast, the risk premium in the counterfactual scenarios ranges between 1.25 and 1.75 percent for the same two bonds. Hence, the risk premium is 30 to 40 percent lower in the counterfactual. Panel (b) analyzes the volatility of excess returns. In the baseline calibration, the volatility of excess returns reaches up to around 6 and 11 percent annually for the 5- and 10-year bond. These values decrease to around 4 and 8 percent in the counterfactual. Panel (c) put average and volatility together and reports the Sharpe ratios. We find that the declines in expected

returns dominate the declines in returns volatilities so that the Sharpe ratios also fall. The Sharpe ratios decrease from around 0.6 to 0.35 for short maturities, a 40 percent drop, but this effect dampens with longer maturities. The lower term premium and Sharpe ratio in turn translate into higher borrowing costs and can distort savings and investment decisions. This exercise put forward a new channel through which improvement in the central bank communication about future decisions can improve their effectiveness.

Figure 6: Counterfactual—Excess Returns

Model-implied unconditional statistics of yields in the baseline calibration (Baseline), in a counterfactual case setting the standard deviation of noise to zero for the expectation component $\sigma_{\eta_1} = 0$ (Perfect signal EH), or the case setting both standard deviation parameters to zero $\sigma_{\eta_1} = \sigma_{\eta_2} = 0$ (Perfect signals). Panel (a): average expected bond excess returns; Panel (b): volatility of bond excess returns; Panel (c): Sharpe ratio.



The lower excess returns and Sharpe ratios are due to the lower variability of the pricing kernel (Equation 14). The reason is that reducing the noise and allowing for perfect information about future short rates lowers the variance of the pricing kernel, which implies a lower covariance of returns with the pricing kernel and, therefore, a lower bond risk premium.

The impact on the yield curve is completely driven by this risk premium channel. In particular, the counterfactual with perfect information leaves almost unchanged the variance of the filtered state variables and that of bond yields. To see why that is the case, Equation (18) can be simplified to:

$$\mathbb{V}\left(\mathcal{P}_{i,t+1|t+1} \middle| \underline{\mathcal{F}_t}\right) = \sigma^2_{\epsilon,i}\left[\phi_i^2 + \mathcal{K}_i\left(1 - \phi_i^2\right)\right].$$

In general, if $\phi_i << 1$, the conditional variances of the factors increase if we increase

the precision of the signal. This arises because the investors' filtering problem produces estimates that are less smooth and track the signal more closely. However, this effect essentially disappears empirically because the yields factors are persistent, i.e., $\phi_i \approx 1$.

4 Robustness of the Variance Decomposition

The stylized facts in Figure 2 rely on the baseline dynamic specification given in Equation (6) as well as the monthly aggregation given in Equation (9). In this section, we confirm that the main results are robust across a range of dynamic specifications and across a different aggregation schemes. We also provide model-free evidence based on survey forecasts data as well as direct regressions.

4.1 Survey-based Evidence

Survey forecasts support the results in Figure 1. To show this, we use quarterly survey forecasts for the 3-month T-bill $SPF_t^{(h)}$ at horizon h from the survey of professional forecasters as a proxy for the expectation component. We then estimate the contemporaneous regressions of forecast changes on the quarterly news components from yields $\Delta\mathcal{P}_{n,t}$:

$$\Delta SPF_{t,t+3}^{(h)} = \alpha_h + \beta_h^\top \sum_{i=1}^{3} \Delta\mathcal{P}_{n,t+i} + \rho_h SPF_t + u_{t,t+3}^{(h)}, \tag{23}$$

where we include a lag of the survey forecasts to account for a potential mean reversion in survey forecasts, which may be due to systematic forecast errors that are correlated with the level of interest rate.

Table 2 reports results for quarterly horizons from 1 to 4 quarters ahead as well as annual horizons of 1 and 2 years. We use the same sample period than in our earlier results. We include the R^2s as well as a partial R^2 measure that corresponds to the share of macro news in the variance of the expectation component (to the extent that the SPF survey is a good measure of expectations):

$$\frac{\mathbb{V}\left(\widehat{\beta}_h^\top \sum_{i=1}^{3} \Delta\mathcal{P}_{n,t+i}\right)}{\mathbb{V}\left(\widehat{\beta}_h^\top \sum_{i=1}^{3} \Delta\mathcal{P}_{n,t+i} + \widehat{\rho}_h SPF_t + \widehat{u}_{t,t+3}^{(h)}\right)}, \tag{24}$$

where $\widehat{\beta}_h$, $\widehat{\rho}_h$ and $\widehat{u}^{(h)}_{t,t+3}$ are the OLS estimates.

Across all quarterly and annual horizons, the R^2s are less than 10 percent and the partial R^2s range between 3 and 8 percent. There is some evidence of a role for the slope component $\Delta\mathcal{P}_{n,2,t}$—the coefficient estimates are close to 0.6 across the forecast horizons—but the estimates are only marginally significant. There is also some evidence for the level component, but only at the shortest horizons and the estimates are again only marginally significant. We conclude that results using survey forecasts are consistent with our earlier results. Revisions in survey forecasts of future short-term interest rates are weakly related to yield changes measured around economic data releases during the same quarter.

Table 2: SPF T-Bill Rate Forecasts and Data Releases

Columns (1)-(6): regressions of quarterly changes in SPF survey forecasts of the average 3-month US T-bill at horizons 1 to 4 quarters ahead and 1 to 2 years (see Equation 23). Columns (7)-(8): regressions survey-based measures of the term premium in the 1-year and 2-year yields, respectively. The partial R^2 measure reports the share of the variance attributed to the news components $\mathcal{P}_n$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

	(1) 1Q	(2) 2Q	(3) 3Q	(4) 4Q	(5) 1Y	(6) 2Y	(7) TP(1Y)	(8) TP(2Y)
lag	−0.029	−0.033	−0.034	−0.033	−0.033	−0.054**	−0.483***	−0.533***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.023)	(0.027)	(0.090)	(0.094)
$\Delta\mathcal{P}_{n,1}$	1.486*	1.438*	1.254	0.955	0.851	0.683	2.635***	2.047**
	(0.813)	(0.831)	(0.801)	(0.750)	(0.900)	(0.967)	(0.833)	(0.910)
$\Delta\mathcal{P}_{n,2}$	0.683**	0.661**	0.588*	0.520*	0.561	0.667*	0.720**	0.192
	(0.320)	(0.327)	(0.315)	(0.295)	(0.354)	(0.381)	(0.333)	(0.358)
$\Delta\mathcal{P}_{n,3}$	0.169	0.108	0.047	−0.019	0.105	−0.236	0.165	0.054
	(0.191)	(0.195)	(0.188)	(0.176)	(0.212)	(0.227)	(0.192)	(0.210)
Constant	0.047	0.057	0.059	0.056	0.034	0.096	0.149***	0.059
	(0.071)	(0.074)	(0.074)	(0.073)	(0.078)	(0.095)	(0.054)	(0.057)
Obs.	87	87	87	87	87	87	87	87
R^2	0.076	0.076	0.074	0.071	0.056	0.105	0.302	0.351
Part. R^2	0.054	0.052	0.05	0.05	0.033	0.083	0.142	0.144

We can also use survey forecasts to provide model-free evidence about the share of the term premium variance due to data releases. To proxy for the term premium, we use the difference between a given yield and the SPF survey forecast of the T-bill rate for an horizon that matches the yield maturity. The last two columns of Table 2 report the results

from regressions that parallel Equation (23) but with this term premium proxy on the left-hand side. The results show that the R^2s are three times higher. The partial R^2s are also twice as high, relative to the results for the expectation components. Consistent with our earlier results, the yields' responses around the data releases are more closely linked to changes in the term premium than in the expectation component.[21]

4.2 Contemporaneous Excess Returns Regressions

Consider the excess returns $rx_{t+h}^{(m)}$ from holding the bond with maturity m for a period of h months, which is given by:

$$rx_{t,t+h}^{(m)} = -\frac{m-h}{h} y_{t+h}^{(m-h)} + \frac{m}{h} y_t^{(m)} - y_t^{(h)}. \tag{25}$$

Then, using Equations (1)-(2) we arrive at the following decomposition:

$$rx_{t,t+h}^{(m)} = \alpha_{h,m} + \beta_{h,m}^\top \mathcal{P}_t + \gamma_{h,m}^\top \sum_{i=1}^{h} \Delta\mathcal{P}_{n,t+i} + \eta_{t,t+h}^{(m)}, \tag{26}$$

where we use the common simplifying assumption that the current components $\mathcal{P}_t$ span the forecasts of the future yields $y_{t+h}^{(m)}$.[22] This representation attributes realized excess returns to three sources: *(i)* the predictable bond risk premium component $\beta_{h,m}^\top \mathcal{P}_t$, *(ii)* the yield changes due to data releases $\gamma_{h,m}^\top \sum_{i=1}^{h} \Delta\mathcal{P}_{n,t+i}$ and *(iii)* a residual component $\eta_{t,t+h}^{(m)}$ due to other information. We can estimate Equation (26) with ordinary least squares and derive an estimator of the share of variance attributed to data releases, given by:

$$\frac{\mathbb{V}\left(\widehat{\gamma}_{h,m}^\top \sum_{i=1}^{h} \Delta\mathcal{P}_{n,t+i}\right)}{\mathbb{V}\left(\widehat{\gamma}_{h,m}^\top \sum_{i=1}^{h} \Delta\mathcal{P}_{n,t+i} + \widehat{\eta}_{t,t+h}^{(m)}\right)}, \tag{27}$$

where $\widehat{\gamma}_{h,m}$ and $\widehat{\eta}_{t,t+h}^{(m)}$ are the OLS estimates.

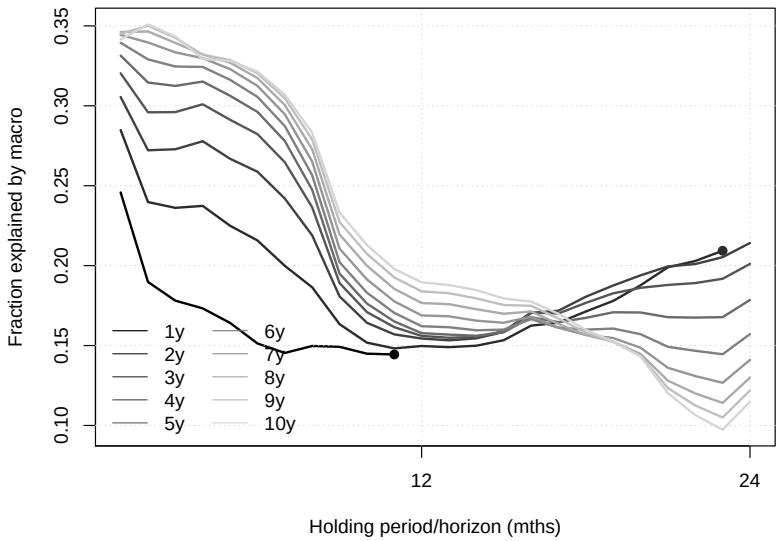
Figure 7 reports this partial R^2 for maturities $m = \{12, 24, \ldots, 120\}$ and for monthly holding periods from one month up to two years. At the one-month horizon, the variance

[21] The results based on survey forecast in Table 2 also suggest that the term premium exhibits reversion to the mean but that the expectation components are more persistent and close to a random walk.

[22] This is a conservative choice. Expanding the information set used to forecast the risk premium component will tend to lower the estimated share of variance attributed to the data releases.

Figure 7: Bond Returns Variance Decomposition

Variance ratios computed using the estimator in Equation (27) based on the excess return regressions in Equation (26). This variance ratio measures the share of the unexpected yield variability over a given horizon that can be attributed to macroeconomic data releases during the same horizon, after controlling for the predictable variations due to a risk premium.



ratio ranges between 25 to 35 percent across the bond maturities. Altavilla, Giannone, and Modugno (2017) provide comparable results for the weekly, monthly and quarterly horizons.[23] At longer horizons, we find that the share of the variance explained by data releases gradually decreases. The overall pattern and level are consistent with our core results.

The regression-based evidence in Figure 7 cannot be extended to longer holding periods. A dynamic model such as the one in Section 2 is needed for that. The main reason is that, as the horizon lengthens, the left-hand side variable in the regression features overlapping observations and the number of distinct observations is small, meaning that reliable inference can be difficult to achieve (e.g., Bauer and Hamilton 2018).

[23] If anything, we find slightly higher ratios at horizons below one quarter, which may be due to the fact that the regression specification in Equation (26) controls for the predictable component of yields.

4.3 Other Model Specifications

The baseline specification in Equation (6) nests the simple VAR(1) as well as more complex models with up to 31 parameters, and potentially includes a moving-average component. We can easily check the results' robustness across a range of models.

Table 3 summarizes the eight cases that we consider. The first group of nested models includes four specifications sharing the restriction that $K_n = K_y$. In these cases, the news components $\mathcal{P}_{n,t}$ drop from the conditional mean dynamics of yields, and we expect that the pattern of decreasing variance ratios in Figure 2 will largely disappear. In the univariate case, it is easy to check that the variance ratio is flat across the horizons. This group of models includes the VAR(1) and the VAR(2) as well as the extension of these two models with one moving-average component. The labels for these two extensions are CM(1) and CM(2,1) The matrix of moving-average coefficients Θ in Equation (6) is restricted to $\Theta = \theta \cdot I$, where θ is a scalar, both for parsimony and to be consistent with the evidence in Feunou and Fontaine (2018).

Table 3: Model Specifications

Taxonomy of model specifications. The number of parameters # in the conditional mean dynamics does not account for the covariance matrix, which has the same number of parameters across specifications.

	K_y	K_n	Θ	#
VAR(1)	0	0	0	12
VAR(2)	K_y	K_y	0	21
CM(1)	0	0	$\theta \cdot I$	13
CM(2,1)	K_y	K_y	$\theta \cdot I$	22
VAR-RR	$\alpha k_y^\top$	$\alpha k_n^\top$	0	20
VAR-UR	K_y	K_n	0	30
CM-RR	$\alpha k_y^\top$	$\alpha k_n^\top$	$\theta \cdot I$	21
CM-UR	K_y	K_n	$\theta \cdot I$	31

A second group of models includes four specifications where $K_n \neq K_y$, which allows data releases to have a distinct impact on the dynamics of yields. Again, the first two cases VAR-RR and VAR-UR have no moving-average component, while the last two cases CM-RR and CM-UR do. The VAR-RR and CM-RR models consider the parsimonious case where only one line linear combination of the news is enough to update the dynamics for

the yield factors (i.e., $K_y = \alpha k_y^\top$ and $K_n = \alpha k_n^\top$). This parsimony is consistent with the evidence in Bauer (2015) and Gurkaynak, Kisacikoglu, and Wright (2018) that the impacts of data releases on yields of different maturities have one dimension.

Table 4 provides diagnostics of these models based on statistical criteria.[24] We report the predictive R^2s for each factor $\mathcal{P}_{i,t}$, the sum of the squared predictive residuals SSR across factors, the log-likelihood as well as the AIC and AIC-c selection criteria. The AIC criteria accounts for the number of parameters and the AIC-c criteria is a refinement that includes an adjustment for the sample size. For every criteria (i.e., in each column), we use the * and $^\blacktriangledown$ superscript to indicate the best and worst performances.

Table 4: Model Selection Criteria

Summary statistics across model specifications. The R^2 measures is $1 - \frac{V[u_t(i)]}{V[\Delta\mathcal{P}_{y,t}(i)]}$ for each term structure factor $\mathcal{P}_{y,t}(i)$, where $\Delta\mathcal{P}_{y,t} = \Delta\mathcal{P}_t - \Delta\mathcal{P}_{n,t}$; SSR is the sum of squared residuals; Lik is the Gaussian log-likelihood; AIC and AICc are the raw and corrected Akaike information criteria, respectively. The symbol * and $^\blacktriangledown$ indicates the best and worst model, respectively.

	$R^2_{\Delta\mathcal{P}_{1,t}}$	$R^2_{\Delta\mathcal{P}_{2,t}}$	$R^2_{\Delta\mathcal{P}_{3,t}}$	SSR	Lik	AIC	AICc
VAR(1)	$2.27^\blacktriangledown$	2.65	$9.35^\blacktriangledown$	$67.10^\blacktriangledown$	$816.6^\blacktriangledown$	−1585.2	−1580.2
VAR(2)	5.75	7.02	9.95	65.90	836.9	−1607.7	−1598.0
CM(1)	2.50	$2.37^\blacktriangledown$	9.49	67.06	817.2	$-1584.5^\blacktriangledown$	$-1579.0^\blacktriangledown$
CM(2,1)	4.15	9.79	10.90	64.96	851.7	−1635.3	$-1624.9^\star$
VAR-RR	5.90	4.08	9.37	66.72	836.4	−1608.9	−1599.7
VAR-UR	$8.29^\star$	9.74	10.16	65.23	846.7	−1609.3	−1593.0
CM-RR	4.12	4.86	10.18	66.21	847.7	−1629.5	−1619.7
CM-UR	6.36	$11.73^\star$	$12.03^\star$	$63.96^\star$	$860.9^\star$	$-1635.7^\star$	−1618.5

The first message from Table 4 is that there is strong evidence against the restriction $K_y = K_n = 0$. One of either the VAR(1) or the CM(1) specifications ranks last across every criteria that we use. Based on the AIC, there is a steep difference of close to 50 points between the our baseline model and the VAR(1) or the CM(1) model. Hence, the increase in the time-series fit is quite unlikely to be due to the higher number of parameters in the more complex models. Instead, the evidence suggests that the improvements are due

[24] Appendix A.4 reports parameter estimates for the eight specifications in Tables-A.3 -A.4 as well as the RMSEs of pricing errors in Table A.5. The RMSEs are very low and essentially the same across models by construction, because every specification uses the same yield factors $\mathcal{P}_t$ in the pricing equation.

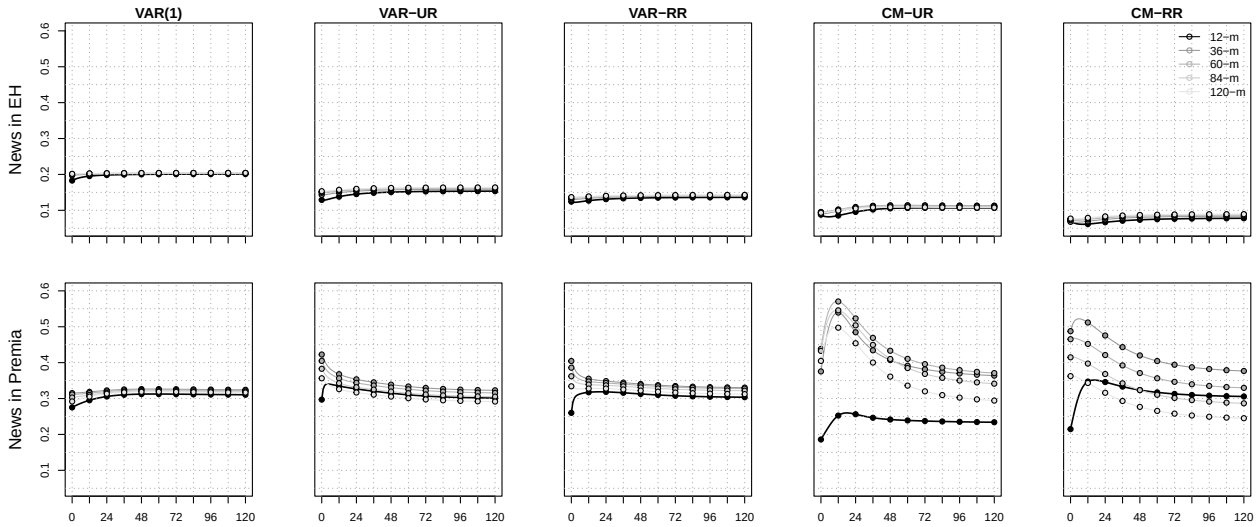
to the larger information set that we use when we distinguish the updates around the economic data releases.

The second message from Table 4 is that there is solid evidence against the restriction $K_n = K_y$ (even if they are both different from zero). There is a single occurrence against this conclusion across all of Table 4, where the AIC-c statistical model criteria favors the CM(2,1) model. Otherwise, the baseline CM-UR specification produces the best performance across model criteria. Based on the likelihood ratio test, the data favor the the CM-UR against CM(2,1) at the 5 percent level (the restriction $K_n = K_y$ is rejected). This statistical difference translates into a predictive R^2 of 6 percent and 12 percent for the level and slope factor, respectively. Again, the evidence supports the use of a richer information set.

Figure 8 reports the variance decompositions for every model that we consider. Our main result is robust. Across all models that allow a distinct role for $\mathcal{P}_{n,t}$, the variance ratios $\mathcal{V}_h^{(m)}$ is higher for the term premium than for the expectation component. Moving from the VAR(1) model to the more flexible models has the main effect of lowering the estimated share attributed to the data releases in the variance ratios of the expectation component.

Figure 8: Share of the Variances Attributed to Data Releases–Robustness

Robustness results. The share of the yields' conditional variance attributed to data releases across model specifications summarized in Table 3. "CM" stands for conditional mean, and indicates the presence of an MA(1) component. "RR" stands for reduced-rank, and indicates that K_n and K_y have rank-one. Monthly data, 1995-2016.



4.4 Other Aggregation Schemes

The baseline aggregation in Equation (9) uses market weights to aggregate the intra-day impacts of data releases up to a monthly time series. The baseline method then uses OLS to map this monthly series of futures price changes into a series for bond yields changes (see Equation 10). We check the robustness of our results across four alternative aggregation schemes.

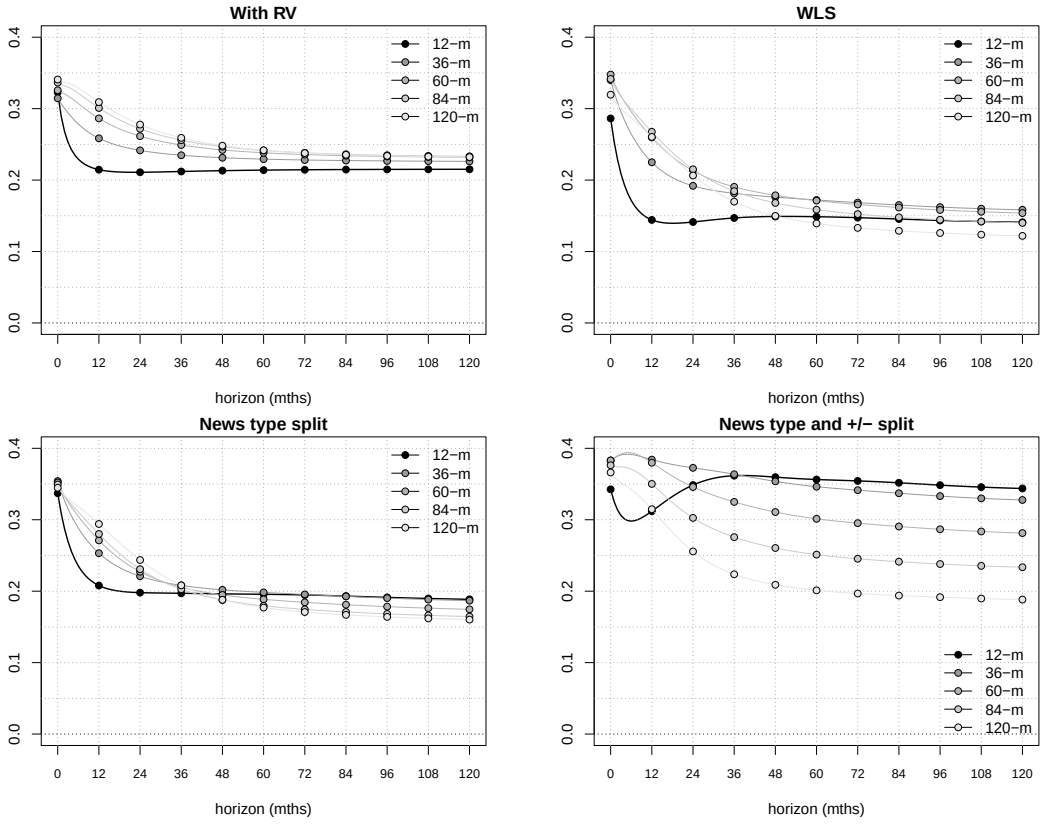
The first alternative is labeled RV and uses information about the realized volatility within the month to aggregate the data. To do this, we estimate the intra-month realized variance-covariance matrix from daily changes in futures prices, and we add all the distinct elements from this matrix to the regression in Equation (10) (this adds the three volatility terms and three correlation terms). In this way, we allow the aggregation to account for a potential correlation between the responses of futures prices and the realized volatility during the month (see Appendix A.6.1). The second alternative is labeled WLS and uses weighted least square (WLS) to estimate Equation (10) to capture potential estimation efficiency gains in the case where the baseline regression residuals exhibit heteroscedasticity (see Appendix A.6.2).

The third alternative uses different weights across the types of data releases in Equation 9. To do this, we first estimate the weights for each news type in daily regressions of the daily yields principal components on the changes in futures prices around news releases. We then use these weights to aggregate the high-frequency measures in the monthly regression (see Appendix A.6.3). Finally, the fourth alternative uses a similar approach, but gives different weights across the types of data releases and depending on the sign of the observed futures price changes (see Appendix A.6.4).

Figure 9 reports the variance decomposition of yields based on these alternative aggregation methods. In all cases, we estimate the dynamics specification given by Equation 6. For every method and virtually all maturities, the variance ratios $\mathcal{V}_h^{(m)}$ starts at a value around 30-40 but declines with the horizon. When we include RV as a control variable, the share attributed to the data releases declines to around 20-25 at the longest horizons. When using weighted least squares, or when using different weights across news types, the share declines to around 15-20 percent. The largest differences relative to our baseline results can be observed for the most flexible aggregation schemes, that is, when we estimate different weights across the types of data releases and the sign of the market impact (bottom-right panel of Figure 9). In this case, the share of the one-year yield's variance

Figure 9: Share of the Variances Attributed to Data Releases–Other Aggregation Schemes

Robustness of the share of the yields' conditional variance attributed to data releases in the baseline CM-UR specification across four aggregation schemes described in Section 4.4 and Appendix A.6.



attributed to news releases is higher, around 35%, and flat across horizons. However, the patterns remain as in the baseline results for bond with longer maturities. Overall, we conclude that the patterns in the baseline results reported in Figure 2 are robust when we used a different method to aggregate the impact of data releases to the monthly frequency.

5 Conclusion

We measure the long-run impact of macroeconomic data releases on bond yields. Macroeconomic news releases are measured using the impact on yields in high-frequency sampling of futures data around the announcements. Our main empirical result shows the impact of these news releases act on bond yields over time; mostly through the term premium. We show theoretically and in a calibration exercise that imperfect information about the response of monetary policy to the new information in data releases can explain our findings. In a counterfactual where the information becomes perfect, the level of the bond risk premium is significantly lower. This significantly adds to the reasons why more work is needed to understand how investors process the arrival of new macroeconomic information and form expectations about future monetary policy.

References

Ai, H. and R. Bansal (2018). Risk preferences and the macroeconomic announcement premium. *Econometrica 86*(4), 1383–1430.

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.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics 73*(2), 251–277.

Andrade, P., R. K. Crump, S. Eusepi, and E. Moench (2016). Fundamental disagreement. *Journal of Monetary Economics 83*, 106 – 128.

Andrei, D. and M. Hasler (2015). Investor attention and stock market volatility. *The review of financial studies 28*(1), 33–72.

Andrei, D., M. Hasler, and A. Jeanneret (2019). Asset pricing with persistence risk. *Review of Financial Studies 32*, 2809–2849.

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*(1-2), 359–403.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The quarterly journal of economics 131*(4), 1593–1636.

Bansal, R. and I. Shaliastovich (2011). Learning and asset-price jumps. *The Review of Financial Studies 24*(8), 2738–2780.

Bauer, M. and J. Hamilton (2018). Robust bond risk premia. *Review of Financial Studies 31*(2), 399–448.

Bauer, M. and E. Swanson (2020). Economic news explains the "Fed information effect". Working paper, Federal Reserve Bank of San Francisco.

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

Beechey, M. J. and J. H. Wright (2009). The high-frequency impact of news on long-term yields and forward rates: Is it real? *Journal of Monetary Economics 56*(4), 535–544.

Berger, D., I. Dew-Becker, and S. Giglio (2020). Uncertainty shocks as second-moment news shocks. *The Review of Economic Studies 87*(1), 40–76.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. De Haan, and D.-J. Jansen (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of economic literature 46*(4), 910–45.

Blinder, A. S., R. Reis, et al. (2005). Understanding the greenspan standard. In *Proceedings-Economic Policy Symposium-Jackson Hole*, Number Aug, pp. 11–96. Federal Reserve Bank of Kansas City.

Brainard, W. C. (1967). Uncertainty and the effectiveness of policy. *The American Economic Review 57*(2), 411–425.

Campbell, J. R., J. D. Fisher, A. Justiniano, and L. Melosi (2017). Forward guidance and macroeconomic outcomes since the financial crisis. *NBER Macroeconomics Annual 31*(1), 283–357.

Campbell, J. Y. and R. J. Shiller (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies 58*, 495–514.

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

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

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.

Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*, 138–160.

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

Dew-Becker, I., S. Giglio, and B. Kelly (2021). Hedging macroeconomic and financial uncertainty and volatility. *Journal of Financial Economics*.

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.

Duffee, G. (2002). Term premia and interest rate forecasts in affine model. *The Journal of Finance 57*, 405–443.

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

Duffee, G. (2021). Macroeconomic news in asset pricing and reality.

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

Fama, E. (1984). Term premiums in bond returns. *Journal of Financial Economics 13*, 529–546.

Faust, J., J. H. Rogers, S.-Y. B. Wang, and J. H. Wright (2007). The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Monetary Economics 54*(4), 1051–1068.

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.

Giacoletti, M., L. Laursen, and K. Singleton (2020). Learning and risk premiums in an arbitrage-free term structure model. *The Journal of Finance*.

Greesnpan, A. (2001). Alan greenspan: transparency in monetary policy. remarks at the Federal Reserve Bank of St. Louis, Economic Policy Conference, St. Louis.

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.

Hamilton, J. D., S. Pruitt, and S. Borger (2011, July). Estimating the market-perceived monetary policy rule. *American Economic Journal: Macroeconomics 3*(3), 1–28.

Hansen, L. and T. Sargent (2020). Macroeconomic uncertainty prices when beliefs are tenuous. *Journal of Econometrics*.

Hanson, S. G., D. O. Lucca, and J. H. Wright (2021). Rate-amplifying demand and the excess sensitivity of long-term rates. *The Quarterly Journal of Economics*.

Hillenbrand, S. (2021). The secular decline in long-term yields around fomc meetings. Working paper, New York University.

Jarociński, M. and P. Karadi (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43.

Johannes, M., L. Lochstoer, and Y. Mou (2016). Learning about consumption dynamics. *The Journal of Finance 71*, 551–600.

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., M. Priebsch, and K. J. Singleton (2014). Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance 69*(3), 1197–1233.

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.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016). A rational theory of mutual funds' attention allocation. *Econometrica 84*(2), 571–626.

Kaminska, I., H. Mumtaz, and R. Sustek (2021). Monetary policy surprises and their transmission through term premia and expected interest rates. Working Paper 914, Bank of England.

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.

Leombroni, M., A. Vedolin, G. Venter, and P. Whelan (2021). Central bank communication and the yield curve. *Journal of Financial Economics*.

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

Ludvigson, S. and S. Ng (2009). Macro factors in bond risk premia. *Review of Financial Studies 22*(12), 5027–5067.

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

Miranda-Agrippino, S. and G. Ricco (2018). The transmission of monetary policy shocks. Discussion paper 13396, CEPR.

Mizrach, B. and C. J. Neely (2008). Information shares in the us treasury market. *Journal of Banking and Finance 32*(7), 1221 – 1233.

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

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics 133*(3), 1283–1330.

Pastor, L. and P. Veronesi (2009). Learning in financial markets. *Annual Review of Financial Economics 1*, 361–381.

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 E. T. Swanson (2012). The bond premium in a dsge model with long-run real and nominal risks. *American Economic Journal: Macroeconomics 4*(1), 105–43.

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.

Savor, P. and M. Wilson (2014). Asset pricing: A tale of two days. *Journal of Financial Economics 113*(2), 171–201.

Steffensen, S., M. Schmeling, and A. Schrimpf (2021). Monetary policy expectation errors. Working paper 175, Dansmark Nationalbank.

Stein, J. C. and A. Sunderam (2018). The fed, the bond market, and gradualism in monetary policy. *The Journal of Finance 73*(3), 1015–1060.

Swanson, E. T. and J. C. Williams (2014). Measuring the effect of the zero lower bound on medium-and longer-term interest rates. *American economic review 104*(10), 3154–85.

Van Binsbergen, J. H., J. Fernández-Villaverde, R. S. Koijen, and J. Rubio-Ramírez (2012). The term structure of interest rates in a dsge model with recursive preferences. *Journal of Monetary Economics 59*(7), 634–648.

Wachter, J. A. and Y. Zhu (2021, 07). A Model of Two Days: Discrete News and Asset Prices. *The Review of Financial Studies*.

A Appendix

A.1 Link with Canonical Macro-finance Models

In the conditional mean equation given by Equation (6), the yield factors $\mathcal{P}_t \in \mathbb{R}^3$ are driven by two sets of uncorrelated shocks: the macro news shocks $\Delta\mathcal{P}_{n,t}$ and other shocks u_t. Hence, we can write:

$$\Delta\mathcal{P}_{t+1} = \mathcal{E}_t + \Delta\mathcal{P}_{n,t+1} + u_{t+1}\,. \tag{A.1}$$

Let us assume that there exists a set of macroeconomic factors $\mathcal{M}_t \in \mathbb{R}^3$ that are relevant to explain yields dynamics. We assume these factors are linear in **the same** present and past shocks that impact the yield factors, such that:

$$\mathcal{M}_{t+1} = \Phi_{\mathcal{M}}\mathcal{M}_t + \Phi_{\mathcal{M},\mathcal{P}}\mathcal{P}_t + \omega_{\mathcal{M}}\Delta\mathcal{P}_{n,t+1} + \omega_{\mathcal{M},\mathcal{P}}u_{t+1}\,. \tag{A.2}$$

In practice, the dimension of $\mathcal{M}_t$ may be less or the same as the dimension of $\mathcal{P}_t$. The dynamics of $\mathcal{P}_t$ and $\mathcal{M}_t$ have the desired VAR(1) representation:

$$\begin{pmatrix} \mathcal{P}_{t+1} \\ \mathcal{M}_{t+1} \end{pmatrix} = \begin{pmatrix} \Phi_{\mathcal{P}} & I \\ \Phi_{\mathcal{M},\mathcal{P}} & \Phi_{\mathcal{M}} \end{pmatrix}\begin{pmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{pmatrix} + \begin{pmatrix} I & I \\ \omega_{\mathcal{M},\mathcal{P}} & \omega_{\mathcal{M}} \end{pmatrix}\begin{pmatrix} u_{t+1} \\ \Delta\mathcal{P}_{n,t+1} \end{pmatrix}, \tag{A.3}$$

where $\mathbb{V}(u_t) = \Omega_u$ and $\mathbb{V}\left(\Delta\mathcal{P}_{n,t}\right) = \Omega_n$. Two identification assumptions are imposed in the above formulation. First, the macro-news shocks $\Delta\mathcal{P}_{n,t}$ enter directly in the yields factors without scaling or rotation. This is merely a consequence of the accounting relationship that the shocks on $\mathcal{P}_t$ are equal to $u_t + \Delta\mathcal{P}_{n,t}$. Second, the top-right of the auto-regressive matrix is equal to identity to set the rotation and the scale of $\mathcal{M}_t$.

Our goal is to map the parameters of Equation (A.3) to the parameters in Equation (6). We can directly see from the previous equation that the conditional mean of $\mathcal{P}_t$ is linear in its past and the past of the macro factors. Thus:

$$\begin{aligned} (\Phi_{\mathcal{P}} - I)\mathcal{P}_t + \mathcal{M}_t &= \mathcal{E}_t \\ \iff (\Phi_{\mathcal{P}} - I)\mathcal{P}_t + \mathcal{M}_t &= K_{\mathcal{P}}\mathcal{P}_t + K_n\Delta\mathcal{P}_{n,t} + K_y\left(\Delta\mathcal{P}_{t-1} - \Delta\mathcal{P}^{(n)}_{t-1}\right) + \Theta\mathcal{E}_{t-1}\,. \end{aligned} \tag{A.4}$$

Then:

$$(\Phi_{\mathcal{P}} - I)\mathcal{P}_t + \mathcal{M}_t = \left(K_{\mathcal{P}} + K_y\right)\mathcal{P}_t + \left(K_n - K_y\right)\Delta\mathcal{P}_{n,t} - K_y\mathcal{P}_{t-1} + \Theta\mathcal{E}_{t-1}\,. \tag{A.5}$$

Since this relationship holds for any date, we directly identify:

$$\Phi_{\mathcal{P}} = I + K_{\mathcal{P}} + K_y\,. \tag{A.6}$$

We rewrite our equality as:

$$\mathcal{M}_t = \left(K_n - K_y\right)\Delta\mathcal{P}_{n,t} - K_y\mathcal{P}_{t-1} + \Theta\mathcal{E}_{t-1}\,. \tag{A.7}$$

Plugging Equation (A.4), we get:

$$\begin{aligned}\mathcal{M}_t &= \left(K_n - K_y\right)\Delta\mathcal{P}_{n,t} - K_y\mathcal{P}_{t-1} + \Theta\left[\left(\Phi_{\mathcal{P}} - I\right)\mathcal{P}_{t-1} + \mathcal{M}_{t-1}\right] \\ &= \left(K_n - K_y\right)\Delta\mathcal{P}_{n,t} - K_y\mathcal{P}_{t-1} + \Theta\left[\left(K_{\mathcal{P}} + K_y\right)\mathcal{P}_{t-1} + \mathcal{M}_{t-1}\right] \\ &= \left(K_n - K_y\right)\Delta\mathcal{P}_{n,t} + \left[\Theta K_{\mathcal{P}} + \left(\Theta - I\right)K_y\right]\mathcal{P}_{t-1} + \Theta\mathcal{M}_{t-1}\,. \end{aligned} \tag{A.8}$$

This formulation provides us with an immediate mapping between the two models:

$$\begin{pmatrix} \mathcal{P}_{t+1} \\ \mathcal{M}_{t+1} \end{pmatrix} = \begin{pmatrix} I + K_{\mathcal{P}} + K_y & I \\ \Theta K_{\mathcal{P}} + \left(\Theta - I\right)K_y & \Theta \end{pmatrix}\begin{pmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{pmatrix} + \begin{pmatrix} I & I \\ 0 & K_n - K_y \end{pmatrix}\begin{pmatrix} u_{t+1} \\ \Delta\mathcal{P}_{n,t+1} \end{pmatrix}. \tag{A.9}$$

A.2 Model properties

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

$$\begin{aligned}\begin{pmatrix} \mathcal{P}_{t+1} \\ \Delta\mathcal{P}_{y,t+1} \end{pmatrix} = &\begin{pmatrix} K_0 \\ K_0 \end{pmatrix} + \begin{pmatrix} I + K_{\mathcal{P}} & K_y + \Theta \\ K_{\mathcal{P}} & K_y + \Theta \end{pmatrix}\begin{pmatrix} \mathcal{P}_t \\ \Delta\mathcal{P}_{y,t} \end{pmatrix} \\ &+ \begin{pmatrix} u_{t+1} \\ \Delta\mathcal{P}_{n,t+1} \end{pmatrix} - \begin{pmatrix} \Theta & -K_n \\ \Theta & -K_n \end{pmatrix}\begin{pmatrix} u_t \\ \Delta\mathcal{P}_{n,t} \end{pmatrix}. \end{aligned} \tag{A.10}$$

We can then easily check the mappings with existing DTSMs. First, 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. Finally, relaxing $\Theta \neq 0$ but with $K_y = K_n = 0$ leads to the family of CM-DTSM models in Feunou and Fontaine (2018). These three yield-only special cases have in common the restriction $K_y = K_n$ and, therefore, that there are 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 $\mathcal{P}_{t+1}$ with coefficients of one for both. However, our use of high-frequency data allows us to identify the differential effect of news and non-news on yields.

There is another reason why this representation is useful: we can easily check conditions for the stationarity and invertibility 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 $\left(\mathcal{P}_t,\, \Delta\mathcal{P}_{y,t+1}\right)$ *is (i) stationary and (ii) invertible if and only if:*

$$(i) \quad \max\left|\text{sp}\begin{pmatrix} I + K_{\mathcal{P}} & K_y + \Theta \\ K_{\mathcal{P}} & K_y + \Theta \end{pmatrix}\right| < 1 \tag{A.11}$$

$$(ii) \quad \max\left|\text{sp}\begin{pmatrix} \Theta & -K_n \\ \Theta & -K_n \end{pmatrix}\begin{pmatrix} 0 & I \\ -I & -I \end{pmatrix}\right| < 1. \tag{A.12}$$

A.3 Construction of the Monthly News Component $\mathcal{P}_{n,t}$

A.3.1 CME 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. Mizrach and Neely (2008) show that for the larger part of our sample, the future contracts were more informative than Treasury securities around data releases. We choose a window of 45 minutes to capture changes in futures prices around each release. 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 illustrates aggregation of high-frequency impacts in Equation (9) for the case of 5-year Treasury futures. Blue bars correspond to change in small windows around macro releases within the month and red bars correspond to change between these windows. Panel (a) reports the results for July 2004. On July 2nd, the 5-year yield declined by 18 bps after the Non-Farm Payroll release (payrolls had increased by 112 thousands month-over-month, less than half the median survey forecast). July 2004 is a typical month when data releases play the larger role: the 5-year yield declined by 22 bps overall, due to the data releases. Panel (b) shows the results for September 1998, in the aftermath of LTCM's financial difficulties, which is an example where data releases played only a small role. Overall, the 5-year yield fell by around 45 basis points during the month but little of it fell in windows with data releases.[25]

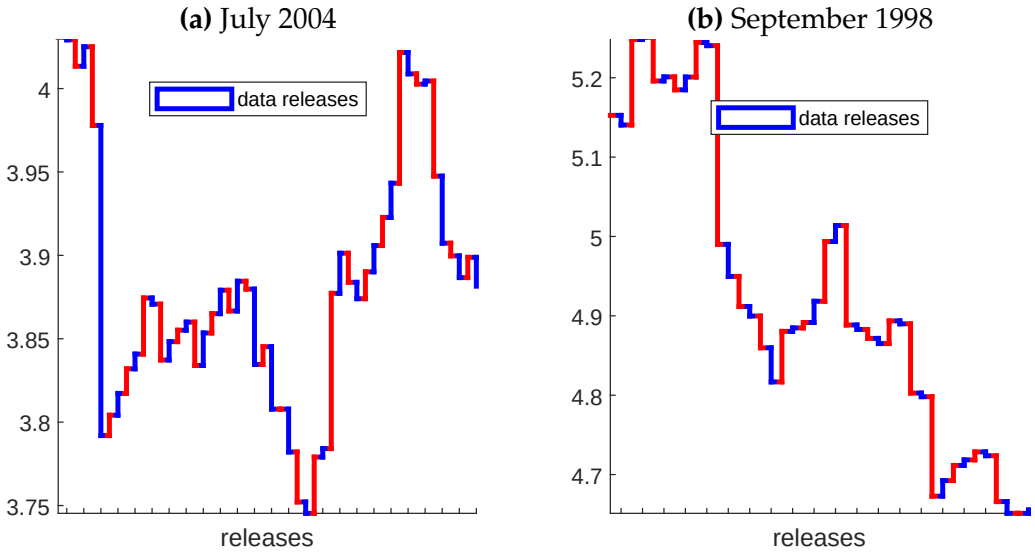
A.3.2 The Baseline Monthly Regression Our approach allows a drift in futures returns to capture its potential premium. For a pre-determined announcement time τ, the conditional mean of futures returns is given by:

$$\mathbb{E}_{\tau-1}\left(\Delta F^{(\mu)}_{t,\tau,j}\right) = a^{(\mu)}_{t,\tau-1,j}\,. \tag{A.13}$$

[25]The largest fall occurred on September 11, 1998, with four other large falls during the month, 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 points), but also partly reflecting a flight to the safety of Treasury bonds (Longstaff 2004; Fontaine and Garcia 2012).

Figure A.1: Daily 5-Year Futures Changes

Changes in the 5-year Treasury futures-implied yields during days with data releases (blue bars) and without data releases (red bars). Panel (a): changes during the month of July 2004; Panel (b): during the month of September 1998. The components $\mathcal{P}_{n,t}$ that we construct sum the effect of all these releases across three maturities for every month in our sample.



Throughout the paper, we assume that this premium is constant through time and across news type, such that:

$$a^{(\mu)}_{t,\tau-1,j} = a^{(\mu)}\,. \tag{A.14}$$

Consider the regression of the monthly changes in one of the principal components $\mathcal{P}^{(i)}_t$ on the total changes in futures prices around data releases during the month t:

$$\Delta\mathcal{P}^{(i)}_t = \beta_i^\top \sum_{j=1}^{J} \sum_{\tau\in[t,t+1)} \left[\Delta F_{t,\tau,j} - \mathbb{E}_{t-1}\left(\Delta F_{t,\tau,j}\right)\right] + \eta^{(i)}_t\,, \tag{A.15}$$

where β_i has length 3. There is no need for an intercept because the left- and right-hand sides have mean zero. Using Equation (A.14), we obtain:

$$\Delta\mathcal{P}^{(i)}_t = \beta_i^\top \left[\sum_{j=1}^{J} \sum_{\tau\in[t,t+1)} \Delta F_{t,\tau,j} - Ja \times \mathcal{J}_t\right] + \eta^{(i)}_t\,, \tag{A.16}$$

where $a = \left(a^{(24)}, a^{(60)}, a^{(120)}\right)^\top$ and $\mathcal{J}_t$ denotes the number of news announced in month t. In principle, it could be possible to identify the vector a, but this is not of interest for our purposes.

Hence, we write in a reduced-form way:

$$\Delta \mathcal{P}_t^{(i)} = \beta_i^\top \left[\sum_{j=1}^{J} \sum_{\tau \in [t,t+1)} \Delta F_{t,\tau,j} \right] - \gamma_i \mathcal{J}_t + \eta_t^{(i)}, \tag{A.17}$$

which leads to Equation (10) in the text.

A.3.3 Additional Results Table A.1 reports regression results for Equation (10). The coefficient estimates are intuitive. The first component $\mathcal{P}_{1,t}$ shows only one significant coefficient estimate at the 10-year maturity, consistent with this component capturing level effects. The second component $\mathcal{P}_{2,t}$ has two significant coefficient estimates with opposite signs at the 2- and 10-year maturities, consistent with this component capturing slope effects across bonds. Finally, $\mathcal{P}_{3,t}$ has three significant coefficient estimates with a V shape across the 2-, 5- and 10-year maturities, consistent with it capturing curvature effects.

Table A.1: Regression Results for Equation (10)

Results from estimating the linear regression of Equation (10) using OLS. The right-hand side variables consist of the 2-year, 5-year, and 10-year futures-implied yields variations around macroeconomic announcements, as defined by Equation (9). $\mathcal{J}_t$ counts the number of news per month.

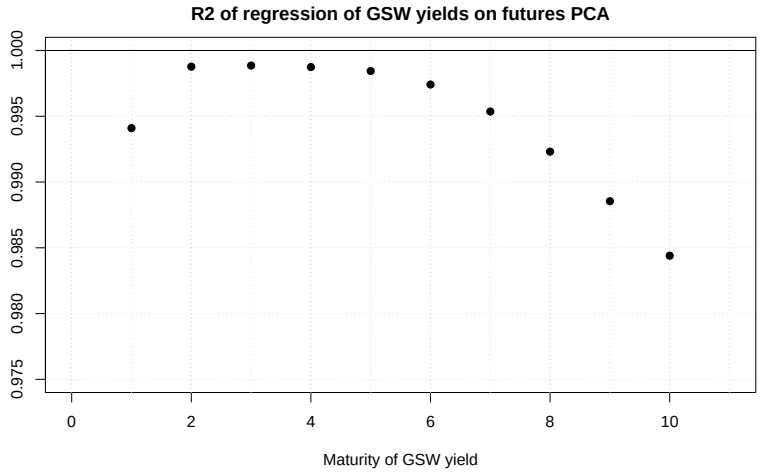
	Dependent variable:		
	$\Delta\mathcal{P}_{1,t}$	$\Delta\mathcal{P}_{2,t}$	$\Delta\mathcal{P}_{3,t}$
$\Delta F_{n,t}^{(2y)}$	0.195	0.949**	1.685***
	(0.155)	(0.376)	(0.638)
$\Delta F_{n,t}^{(5y)}$	−0.241	1.230*	−7.038***
	(0.302)	(0.734)	(1.247)
$\Delta F_{n,t}^{(10y)}$	0.806***	−3.751***	5.383***
	(0.279)	(0.678)	(1.151)
$\mathcal{J}_t$	−0.0005	0.0005	−0.00004
	(0.0004)	(0.001)	(0.002)
Observations	263	263	263
R^2	0.331	0.278	0.176
Note:		*p<0.1; **p<0.05; ***p<0.01	

A.3.4 Comparing Yields and Futures Data 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 on the three principal components of futures yields, using monthly data. Figure A.2 shows that the R^2 in regressions of bond yields on futures yields range between 0.985 and 0.999.

Second, Figure A.3 shows that the first two principals components from Treasury yields and futures (i.e., level and slope) are essentially identical in the time series, while the third principal

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

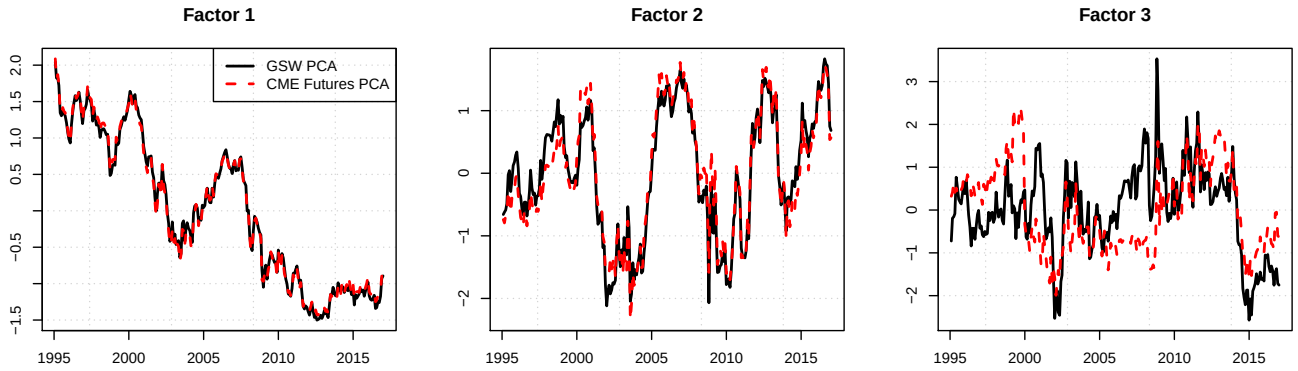
R^2 in regressions of GSW zero-coupon yield $y_t^{(m)}$ on F_t: three futures-implied yields for bonds with 2, 5 and 10 years.



component (curvature) exhibits only small differences.

Figure A.3: Yield and Futures PCA–Time Series

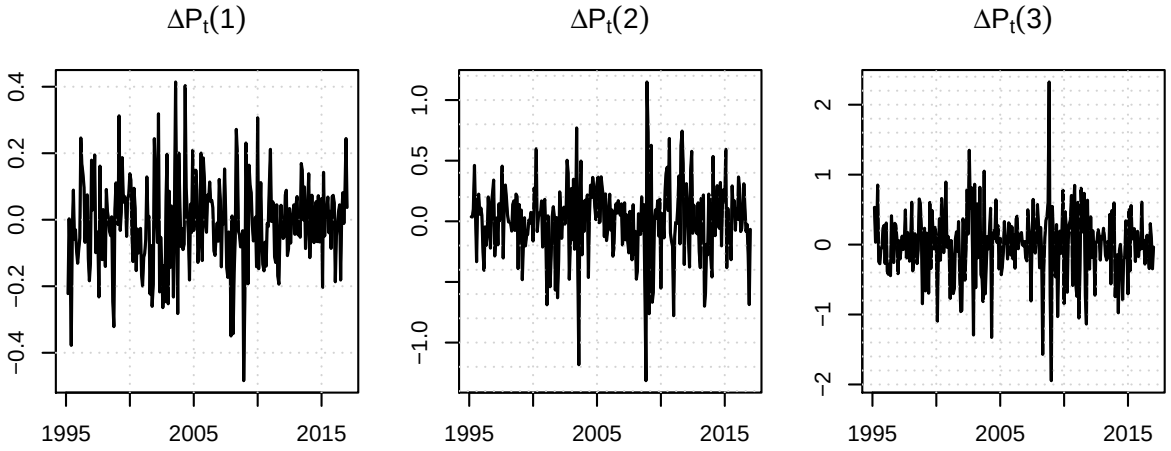
Times-series of bond yield PCAs and futures yield PCA.



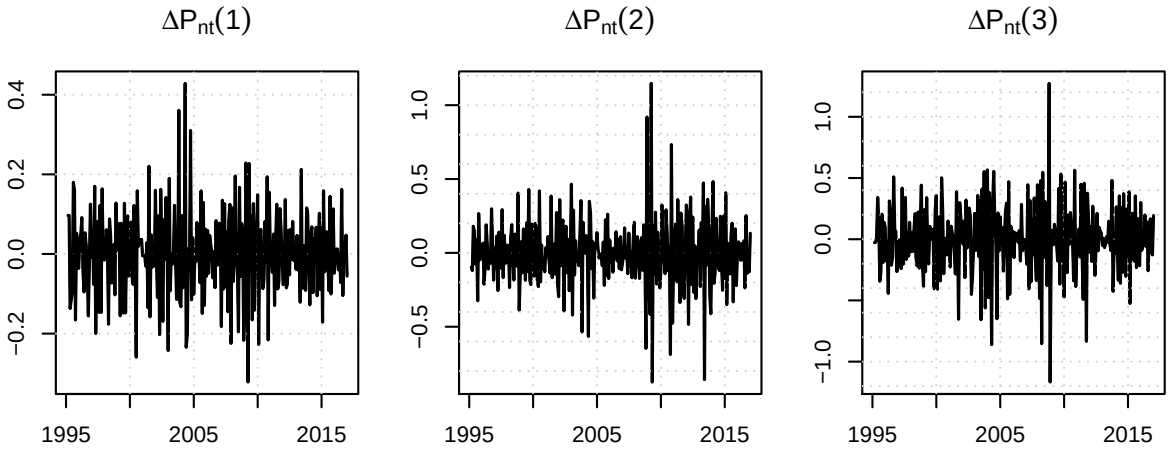
A.3.5 Auto-correlations Figure A.4 shows the time-series of monthly changes in futures $\Delta\mathcal{P}_t$ and its news component $\Delta\mathcal{P}_{n,t}$. The patterns clearly suggest that $\Delta\mathcal{P}_t$ has a slowly-moving component but that $\mathcal{P}_{n,t}$ is white noise. Inspection 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 $\mathcal{P}_{n,t}$ is white noise (unreported).

Figure A.4: Monthly Changes $\Delta\mathcal{P}_t$ and $\Delta\mathcal{P}_{n,t}$

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



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



A.3.6 Principal component analysis of $\Delta\mathcal{P}_t$ Table A.2 shows that the first principal component of $\Delta\mathcal{P}_t$, $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$ explains a very similar 95 percent of the total variance in each case.

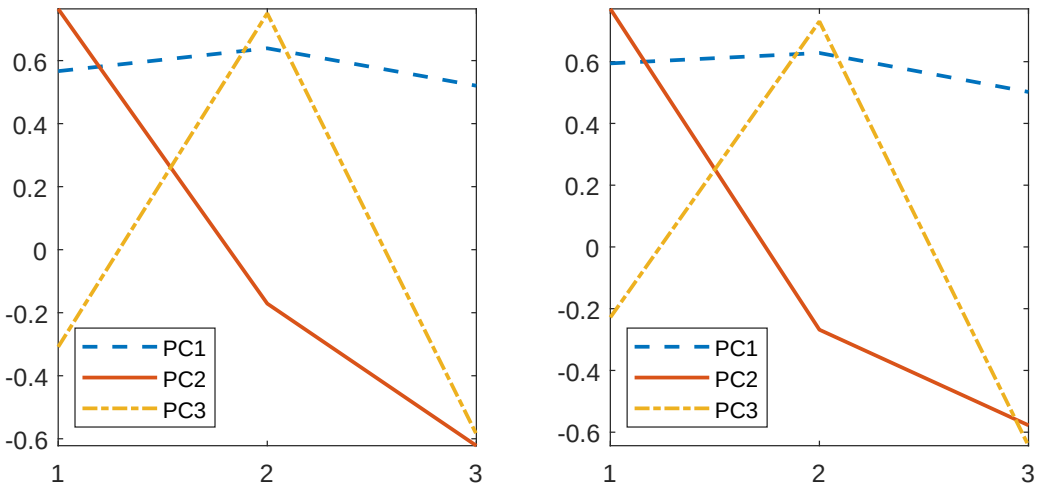
Table A.2: Share of variance explained by three principal components.

	$\Delta\mathcal{P}_t$	$\Delta\mathcal{P}_{n,t}$	$\Delta\mathcal{P}_{y,t}$
PC1	94.36	95.05	93.33
PC2	4.94	4.32	5.60
PC3	0.71	0.64	1.07

Second, Figure A.5 shows that in every case, the loadings have the usual interpretation in terms of level, slope and curvature despite being extracted from yields changes. We find that the level component of $\Delta\mathcal{P}_t$ inherits its predictable dynamics (see Figure A.4). By contrast, the news components $\Delta\mathcal{P}_{n,t}$ appear to be largely unpredictable.

Figure A.5: Principal Component Analysis
PCA weights extracted from $\Delta\mathcal{P}_t$ and $\Delta\mathcal{P}_{n,t}$, respectively.

(a) Monthly Changes **(b)** News Components



The takeaways are that the monthly yield changes, the monthly news components and the monthly non-news components each exhibits one large leading factor. To be clear, the almost identical factor structure in Table A.2 does not imply that the monthly components contain the same information. Instead, this table says the changes of yields at different maturities can be summarized by one factor—whether around news or between news—and that this factor acts as a level.

This is consistent with Bauer (2015), who finds that the effect of macroeconomic news on yields of different maturities can be summarized by one factor. Gurkaynak et al. (2018) also find a strong common response of interest rates to data releases. This is an important feature of the data that we can use to maintain parsimony in term structure models: the information in $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$ can be summarized with one linear combination.

A.4 Estimation

The conditional likelihood of $\mathcal{P}_{t+1}$ is available in closed-form based on Equation (6), which can be estimated by updating $\mathcal{E}_t$ recursively to recover the forecast errors:

$$\Delta\mathcal{P}_{t+1} - \mathbb{E}\left(\Delta\mathcal{P}_{t+1}|\underline{\mathcal{F}_t}\right) = \Delta\mathcal{P}_{n,t+1} + u_{t+1}. \tag{A.18}$$

We impose that risk-neutral and historical conditional covariances are the same, given by $\Sigma = \Omega_u + \Omega_n$. The estimation of dynamic parameters proceeds by quasi-maximum likelihood. To compute the likelihood, we use the plug-in estimators of the variances $\Omega_u = \mathbb{V}\left(u_{t+1}|\underline{\mathcal{F}_t}\right)$ and $\Omega_n = \mathbb{V}\left(\Delta\mathcal{P}_{n,t+1}|\underline{\mathcal{F}_t}\right)$ and we assume $Cov\left(u_{t+1},\, \Delta\mathcal{P}_{n,t+1}|\underline{\mathcal{F}_t}\right) = 0$. Indeed, based on the regression in Equation (10), we have that $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_t - \Delta\mathcal{P}_{n,t}$ are orthogonal. Using the law of total variance, we have that $\mathbb{E}\left[Cov\left(u_{t+1},\, \Delta\mathcal{P}_{n,t+1}|\underline{\mathcal{F}_t}\right)\right] + Cov\left[\mathbb{E}\left(u_{t+1}|\underline{\mathcal{F}_t}\right),\, \mathbb{E}\left(\Delta\mathcal{P}_{n,t+1}|\underline{\mathcal{F}_t}\right)\right] = 0$. The second term is null, since $\Delta\mathcal{P}_{n,t}$ is not predictable. If the conditional covariance is constant, then we get the desired result. Finally, Appendix A.2 shows that stationarity and invertibility properties can be verified based on simple parameter computations.

Table A.3: Parameter Estimates—Variance

This table reports the variance parameter estimates for the specifications described in Table 3. The CM-UR model is our baseline.

	VAR(1)	VAR(2)	VAR-UR	VAR-RR	CM(1)	CM(2,1)	CM-UR	CM-RR
$\sqrt{\Omega_u}(1,1)$	0.1061	0.1042	0.1028	0.1041	0.106	0.1051	0.1039	0.1051
$\sqrt{\Omega_u}(2,1)$	−0.1359	−0.1391	−0.1366	−0.1446	−0.136	−0.1399	−0.1366	−0.1432
$\sqrt{\Omega_u}(3,1)$	−0.2055	−0.2103	−0.2146	−0.2083	−0.2069	−0.2185	−0.2206	−0.2181
$\sqrt{\Omega_u}(2,2)$	0.2184	0.2094	0.2066	0.2105	0.2188	0.2043	0.2033	0.2102
$\sqrt{\Omega_u}(3,2)$	−0.1169	−0.1205	−0.1211	−0.1309	−0.1164	−0.135	−0.1378	−0.1507
$\sqrt{\Omega_u}(3,3)$	0.3489	0.3431	0.3397	0.3422	0.3479	0.3297	0.3237	0.3252
$\sqrt{\Sigma}(1,1)$	0.13	0.1284	0.1273	0.1284	0.1299	0.1291	0.1282	0.1292
$\sqrt{\Sigma}(2,1)$	−0.1856	−0.1884	−0.1866	−0.1929	−0.1857	−0.189	−0.1865	−0.1917
$\sqrt{\Sigma}(3,1)$	−0.2352	−0.2389	−0.2421	−0.2372	−0.2364	−0.2456	−0.2472	−0.2453
$\sqrt{\Sigma}(2,2)$	0.2405	0.232	0.2295	0.2326	0.2409	0.2274	0.2266	0.2325
$\sqrt{\Sigma}(3,2)$	−0.1364	−0.1402	−0.1412	−0.1493	−0.136	−0.1538	−0.1566	−0.1677
$\sqrt{\Sigma}(3,3)$	0.3815	0.3763	0.3733	0.3754	0.3806	0.364	0.3586	0.3599

Table A.4: Parameter Estimates—Conditional Mean

This table reports the parameter estimates for the specifications described in Table 3. |AR| and |MA| provide the maximal modulus of eigenvalues of the autoregressive and moving average matrix. The CM-UR model is our baseline.

	VAR(1)	VAR(2)	VAR-UR	VAR-RR	CM(1)	CM(2,1)	CM-UR	CM-RR
$K_0(1)$	−0.0113	−0.01	−0.0089	−0.0088	−0.0137	−0.0024	−0.0029	−0.0024
$K_0(2)$	0.0052	0.0103	0.0095	0.009	0.0062	0.0051	0.0059	0.0027
$K_0(3)$	−0.0031	−0.0046	−0.0038	−0.0038	−0.0035	$-3{\cdot}10^{-4}$	$6{\cdot}10^{-4}$	0.0034
$K_z(1,1)$	−0.0118	−0.0112	−0.0129	−0.0116	−0.0141	−0.0023	−0.0047	−0.003
$K_z(2,1)$	$3{\cdot}10^{-4}$	−0.0011	0.0045	$7{\cdot}10^{-4}$	$3{\cdot}10^{-4}$	−0.0022	0.0017	$-7{\cdot}10^{-4}$
$K_z(3,1)$	0.0603	0.0607	0.0612	0.0602	0.0727	0.0171	0.0256	0.0205
$K_z(1,2)$	0.0078	0.0066	0.0067	0.0067	0.0095	$9{\cdot}10^{-4}$	0.0017	0.0016
$K_z(2,2)$	−0.0373	−0.0416	−0.041	−0.039	−0.0429	−0.0144	−0.0171	−0.0138
$K_z(3,2)$	0.019	0.0232	0.0246	0.0194	0.021	0.0029	0.0074	0.003
$K_z(1,3)$	−0.008	$1{\cdot}10^{-4}$	0.0044	$-5{\cdot}10^{-4}$	−0.0107	0.001	0.0037	0.0012
$K_z(2,3)$	−0.0204	−0.0176	−0.0275	−0.009	−0.02	−0.0062	−0.0136	−0.0081
$K_z(3,3)$	−0.1201	−0.1177	−0.1145	−0.1223	−0.1406	−0.0189	−0.0291	−0.0226
$K_u(1,1)$	0	0.1554	0.251	0.2526	0	0.0456	0.0932	0.098
$K_u(2,1)$	0	0.4751	0.3885	0.3848	0	0.3294	0.3457	0.1385
$K_u(3,1)$	0	−0.1535	−0.0694	−0.0734	0	0.1646	0.2554	0.4854
$K_u(1,2)$	0	0.062	0.0657	0.0631	0	0.0342	0.0477	0.0256
$K_u(2,2)$	0	0.1344	0.0998	0.096	0	0.0184	−0.0071	0.0362
$K_u(3,2)$	0	−0.0875	−0.0529	−0.0183	0	0.0972	0.0905	0.1268
$K_u(1,3)$	0	−0.0142	$-9{\cdot}10^{-4}$	0.0123	0	−0.0189	−0.0085	−0.0116
$K_u(2,3)$	0	0.1242	0.0837	0.0188	0	0.0819	0.0607	−0.0164
$K_u(3,3)$	0	−0.0725	−0.0677	−0.0036	0	−0.1526	−0.1794	−0.0574
$K_n(1,1)$	0	0.1554	−0.1793	0.0587	0	0.0456	−0.1849	−0.0191
$K_n(2,1)$	0	0.4751	1.0692	0.0894	0	0.3294	0.8071	−0.027
$K_n(3,1)$	0	−0.1535	−0.6344	−0.0171	0	0.1646	−0.4359	−0.0947
$K_n(1,2)$	0	0.062	0.008	0.0508	0	0.0342	−0.0255	0.0094
$K_n(2,2)$	0	0.1344	0.3135	0.0775	0	0.0184	0.2176	0.0132
$K_n(3,2)$	0	−0.0875	−0.2939	−0.0148	0	0.0972	−0.0546	0.0464
$K_n(1,3)$	0	−0.0142	−0.0992	0.0058	0	−0.0189	−0.0894	−0.0108
$K_n(2,3)$	0	0.1242	0.3874	0.0088	0	0.0819	0.2864	−0.0152
$K_n(3,3)$	0	−0.0725	−0.1238	−0.0017	0	−0.1526	−0.0856	−0.0534
θ	0	0	0	0	−0.1934	0.7447	0.6614	0.6815
|AR|	0.98	0.982	0.98	0.981	0.979	0.991	0.987	0.985
|MA|	0	0	0	0	0.19	0.75	0.66	0.68

Table A.5: RMSE of Pricing Errors (in bps)

Pricing errors are produced after estimating the risk-neutral dynamics associated with every model. We minimize the equally-weighted sum of squared residuals of each yield maturity to estimate $K_0^{\mathbb{Q}}$ and $K^{\mathbb{Q}}$ associated with the risk-neutral dynamics $\mathcal{P}_t = K_0^{\mathbb{Q}} + K^{\mathbb{Q}}\mathcal{P}_{t-1} + \sqrt{\Sigma}\varepsilon_t^{\mathbb{Q}}$, where $\varepsilon_t^{\mathbb{Q}}$ is a normalized Gaussian white noise vector under the risk-neutral measure. Σ is inherited from the maximum likelihood estimation of the physical measure.

Maturity	12m	36m	60m	84m	120m
VAR(1)	1.97	1.82	1.45	1.62	2.33
VAR(2)	1.97	1.82	1.45	1.62	2.33
VAR-UR	1.97	1.82	1.45	1.62	2.33
VAR-RR	1.97	1.82	1.45	1.62	2.33
CM(1)	1.97	1.82	1.45	1.62	2.33
CM(2,1)	1.97	1.82	1.45	1.62	2.33
CM-UR	1.97	1.82	1.45	1.62	2.33
CM-RR MA	1.97	1.82	1.45	1.62	2.33

A.5 A Two-Factor Bond Pricing Model with Learning

A.5.1 The Investor's Kalman Filter Suppose $\mathcal{P}_t$ is a scalar (to simplify the notation). Then, the one-factor state-space model can be expressed as:

$$\begin{pmatrix} \mathcal{P}_t \\ \mathcal{P}_{t-1} \\ z_t \end{pmatrix} \Big| \underline{\mathcal{F}_{t-1}} \sim \mathcal{N}\left[\begin{pmatrix} \mu + \phi\mathcal{P}_{t-1|t-1} \\ \mathcal{P}_{t-1|t-1} \\ \mu + \phi\mathcal{P}_{t-1|t-1} \end{pmatrix}, \begin{pmatrix} \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 & \phi\omega_{t-1|t-1} & \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 \\ \phi\omega_{t-1|t-1} & \omega_{t-1|t-1} & \phi\omega_{t-1|t-1} \\ \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 & \phi\omega_{t-1|t-1} & \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 + \sigma_\eta^2 \end{pmatrix} \right] \tag{A.19}$$

where $\omega_{t-1|t-1} = \mathbb{V}\left[\mathcal{P}_{t-1}|\underline{F_{t-1}}\right]$. The case with two factors is a simple vector extension where the ϕ and $\mathcal{K}$ are a diagonal matrix. The Kalman gain is easily obtained as:

$$\begin{pmatrix} \phi\omega_{t-1|t-1} \\ \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 \end{pmatrix}' \begin{pmatrix} \omega_{t-1|t-1} & \phi\omega_{t-1|t-1} \\ \phi\omega_{t-1|t-1} & \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 + \sigma_\eta^2 \end{pmatrix}^{-1} = \begin{pmatrix} \phi(1-\mathcal{K}) \\ \mathcal{K} \end{pmatrix}' \tag{A.20}$$

where $\mathcal{K} = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}$. Thus, the update $\mathcal{P}_{t|t}$ is given by:

$$\begin{aligned} \mathcal{P}_{t|t} &= \mu + \phi\mathcal{P}_{t-1|t-1} + \begin{pmatrix} \phi(1-\mathcal{K}) \\ \mathcal{K} \end{pmatrix}' \begin{pmatrix} \mathcal{P}_{t-1} - \mathcal{P}_{t-1|t-1} \\ \eta_t + \varepsilon_t + \phi\left(\mathcal{P}_{t-1} - \mathcal{P}_{t-1|t-1}\right) \end{pmatrix} \\ &= \mu + \phi\mathcal{P}_{t-1} + \mathcal{K}\left(\varepsilon_t + \eta_t\right). \end{aligned} \tag{A.21}$$

Using the Kalman filter recursions, the filtering error variance is then given by:

$$\begin{aligned}\omega_{t|t} &= \phi^2\omega_{t-1|t-1} - \begin{pmatrix} \phi(1-\mathcal{K}) \\ \mathcal{K} \end{pmatrix}' \begin{pmatrix} \phi\omega_{t-1|t-1} \\ \phi^2\omega_{t-1|t-1} + \sigma_\varepsilon^2 \end{pmatrix} \\ &= (1-\mathcal{K})\sigma_\varepsilon^2 \,. \end{aligned} \tag{A.22}$$

Therefore, the complete state-space model is given by:

$$\begin{aligned}\begin{pmatrix} \mathcal{P}_t \\ \mathcal{P}_{t-1} \\ z_t \end{pmatrix} \Big| \underline{\mathcal{F}_{t-1}} \;\sim\; \mathcal{N}\Bigg[&\begin{pmatrix} \mu + \phi\mathcal{P}_{t-1|t-1} \\ \mathcal{P}_{t-1|t-1} \\ \mu + \phi\mathcal{P}_{t-1|t-1} \end{pmatrix}, \\ &\begin{pmatrix} \sigma_\varepsilon^2\left[1+\phi^2(1-\mathcal{K})\right] & \sigma_\varepsilon^2\phi(1-\mathcal{K}) & \sigma_\varepsilon^2\left[1+\phi^2(1-\mathcal{K})\right] \\ \sigma_\varepsilon^2\phi(1-\mathcal{K}) & (1-\mathcal{K})\sigma_\varepsilon^2 & \sigma_\varepsilon^2\phi(1-\mathcal{K}) \\ \sigma_\varepsilon^2\left[1+\phi^2(1-\mathcal{K})\right] & \sigma_\varepsilon^2\phi(1-\mathcal{K}) & \sigma_\varepsilon^2\left[1+\phi^2(1-\mathcal{K})\right] + \sigma_\eta^2 \end{pmatrix} \Bigg] .\end{aligned} \tag{A.23}$$

A.5.2 Bond Pricing The conditional mean of the filtered factors under the risk-neutral measure is given by:

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\begin{pmatrix} \mathcal{P}_{1,t+1|t+1} \\ \mathcal{P}_{2,t+1|t+1} \end{pmatrix} \Big| \underline{\mathcal{F}_t} \right] &= \begin{pmatrix} \mu_1 + \lambda_{0,1} \times \sigma_{\varepsilon,1}^2\left[1+\phi_1^2(1-\mathcal{K}_1)\right] \\ \mu_2 + \lambda_{0,2} \times \sigma_{\varepsilon,2}^2\left[1+\phi_2^2(1-\mathcal{K}_2)\right] \end{pmatrix} \\ + \quad &\begin{pmatrix} \phi_1 + \sigma_{\varepsilon,1}^2\left[1+\phi_1^2(1-\mathcal{K}_1)\right]\Lambda_1 & \sigma_{\varepsilon,1}^2\left[1+\phi_1^2(1-\mathcal{K}_1)\right]\Lambda_{1,2} \\ 0 & \phi_2 + \sigma_{\varepsilon,2}^2\left[1+\phi_2^2(1-\mathcal{K}_2)\right]\Lambda_2 \end{pmatrix} \begin{pmatrix} \mathcal{P}_{1,t|t} \\ \mathcal{P}_{2,t|t} \end{pmatrix} \\ &= \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} \begin{pmatrix} \mathcal{P}_{1,t|t} \\ \mathcal{P}_{2,t|t} \end{pmatrix}, \end{aligned}$$

where we defined $\mu^{\mathbb{Q}}$ and $\Phi^{\mathbb{Q}}$ implicitly. Therefore, if the short rate is given by $r_t = \alpha + \beta\mathcal{P}_{1,t|t}$ then the yields are linear combinations of both factors:

$$y_t^{(m)} = A_m + B_m'\mathcal{P}_{t|t} = A_m + B_{1,m}\mathcal{P}_{1,t|t} + B_{2,m}\mathcal{P}_{2,t|t} \,, \tag{A.24}$$

where $A_m = -\mathcal{A}_m/m$, $B_m = -\mathcal{B}_m/m$, and

$$\begin{aligned}
\mathcal{A}_m &= -\alpha + \mathcal{A}_{m-1} + \mathcal{B}'_{m-1}\mu^{\mathbb{Q}} + \tfrac{1}{2}\mathcal{B}^2_{1,m-1}\sigma^2_{\varepsilon,1}\left[1+\phi_1^2(1-\mathcal{K}_1)\right] + \tfrac{1}{2}\mathcal{B}^2_{2,m-1}\sigma^2_{\varepsilon,2}\left[1+\phi_2^2(1-\mathcal{K}_2)\right] \\
\mathcal{B}_m &= -(\beta_1,\,0)' + \Phi^{\mathbb{Q}'}\mathcal{B}_{m-1} = -\left(I-\Phi^{\mathbb{Q}'}\right)^{-1}\left(I-\Phi^{\mathbb{Q}'^m}\right)\begin{pmatrix}\beta\\0\end{pmatrix} \\
&= \begin{pmatrix} -\frac{1-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}}}\beta \\ \frac{\phi_{1,2}^{\mathbb{Q}}}{1-\phi_2^{\mathbb{Q}}}\left[\frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{\phi_2^{\mathbb{Q}}-\phi_1^{\mathbb{Q}}} - \frac{1-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}}}\right]\beta \end{pmatrix}.
\end{aligned} \tag{A.25}$$

A.5.3 Variance Decomposition of a Single Factor Iterating forward the filtered factor dynamics, we have:

$$\mathcal{P}_{t+h|t+h} = \frac{1-\phi^h}{1-\phi}\mu + \phi^h\mathcal{P}_t + \mathcal{K}\left(\varepsilon_{t+h}+\eta_{t+h}\right) + \sum_{j=1}^{h-1}\phi^j\varepsilon_{t+h-j}\,, \tag{A.26}$$

where ε_t is not observable in real time for the investor. Therefore, we decompose the shocks in its two orthogonal components:

$$\begin{aligned}
\varepsilon_{t+h-j} &= \frac{Cov\left(\varepsilon_{t+h-j},\,\mathcal{K}(\varepsilon_{t+h-j}+\eta_{t+h-j})\,\middle|\,\underline{\mathcal{F}_t}\right)}{\mathbb{V}\left(\mathcal{K}(\varepsilon_{t+h-j}+\eta_{t+h-j})\,\middle|\,\underline{\mathcal{F}_t}\right)}\mathcal{K}(\varepsilon_{t+h-j}+\eta_{t+h-j}) + \varepsilon^{\perp}_{t+h-j} \\
&= \frac{\sigma^2_\varepsilon}{\sigma^2_\varepsilon+\sigma^2_\eta}(\varepsilon_{t+h-j}+\eta_{t+h-j}) + \varepsilon^{\perp}_{t+h-j} \\
&= \mathcal{K}(\varepsilon_{t+h-j}+\eta_{t+h-j}) + \varepsilon^{\perp}_{t+h-j}.
\end{aligned} \tag{A.27}$$

Denoting by $\Sigma = \sigma^2_\varepsilon + \sigma^2_\eta$, we have that $\sigma^2_\varepsilon = \mathcal{K}\Sigma$, $\sigma^2_\eta = (1-\mathcal{K})\Sigma$ and:

$$\mathbb{V}\left(\varepsilon^{\perp}_{t+h-j}\,|\,\underline{\mathcal{F}}_t\right) = \mathcal{K}\Sigma - \mathcal{K}^2\Sigma = \mathcal{K}(1-\mathcal{K})\Sigma\,. \tag{A.28}$$

This orthogonal component $\varepsilon^{\perp}_t$ is the part of the fundamental shock "missed" by the representative investor when performing bayesian updating. Thus, it is equivalent to the filtering error of the investor. We can be convinced of this by looking at the formula for the filtering error itself:

$$\mathcal{P}_t - \mathcal{P}_{t|t} = (1-\mathcal{K})\varepsilon_t - \mathcal{K}\eta_t\,. \tag{A.29}$$

Its conditional variance is given by:

$$\mathbb{V}\left(\mathcal{P}_t - \mathcal{P}_{t|t}\,|\,\underline{\mathcal{F}_t}\right) = (1-\mathcal{K})^2\mathcal{K}\Sigma + \mathcal{K}^2(1-\mathcal{K})\Sigma = \mathcal{K}(1-\mathcal{K})\Sigma\,. \tag{A.30}$$

Notice that the conditioning in the variance is irrelevant since, by construction, the filtering errors are orthogonal to the information set spanned by $\underline{\mathcal{F}_t}$. We can now easily perform the variance

decomposition of the factor:

$$\begin{aligned}
\mathbb{V}\left(\mathcal{P}_{t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right) &= \phi^{2h}\mathbb{V}\left(\mathcal{P}_t - \mathcal{P}_{t|t} \,|\, \underline{\mathcal{F}_t}\right) + \mathcal{K}^2\Sigma + \sum_{j=1}^{h-1}\phi^{2j}\left[\mathcal{K}^2\Sigma + \mathcal{K}(1-\mathcal{K})\Sigma\right] \\
&= \mathcal{K}^2\Sigma\left(1 + \frac{\phi^2 - \phi^{2h}}{1-\phi^2}\right) + \mathcal{K}(1-\mathcal{K})\Sigma\left(\phi^{2h} + \frac{\phi^2 - \phi^{2h}}{1-\phi^2}\right) \\
&= \mathcal{K}^2\Sigma\left(\frac{1-\phi^{2h}}{1-\phi^2}\right) + \mathcal{K}(1-\mathcal{K})\Sigma\left(\phi^2\frac{1-\phi^{2h}}{1-\phi^2}\right). \qquad \text{(A.31)}
\end{aligned}$$

The first bracket is the component related to data releases, while the second part is related to filtering errors. Looking at the ratio of the first component over the total variance, we obtain the result:

$$\mathcal{V} := \frac{\mathcal{K}}{\mathcal{K} + (1-\mathcal{K})\phi^2}. \qquad \text{(A.32)}$$

A.5.4 Proof of Proposition 1 The variance of the yield $y_t^{(m)}$ is given by:

$$\mathbb{V}\left(y_{t+h}^{(m)} \,|\, \underline{\mathcal{F}_t}\right) = B_{1,m}^2\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right) + B_{2,m}^2\mathbb{V}\left(\mathcal{P}_{2,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right). \qquad \text{(A.33)}$$

Denoting by $\mathcal{V}_i$ the share of the variance of factor i attributable to macro-news, we can write the variance decomposition of any yield as:

$$\begin{aligned}
\mathcal{V}_h^{(m)} &= \frac{B_{1,m}^2\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)\mathcal{V}_{1,h} + B_{2,m}^2\mathbb{V}\left(\mathcal{P}_{2,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)\mathcal{V}_{2,h}}{B_{1,m}^2\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right) + B_{2,m}^2\mathbb{V}\left(\mathcal{P}_{2,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)} \\
&= \omega_h^{(m)}\mathcal{V}_1 + \left(1-\omega_h^{(m)}\right)\mathcal{V}_2, \qquad \text{(A.34)}
\end{aligned}$$

where

$$\begin{aligned}
\omega_h^{(m)} &= \frac{B_{1,m}^2\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)}{B_{1,m}^2\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right) + B_{2,m}^2\mathbb{V}\left(\mathcal{P}_{2,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)} \\
&= \frac{1}{1 + \frac{B_{2,m}^2}{B_{1,m}^2} \times \frac{\mathbb{V}\left(\mathcal{P}_{2,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)}{\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right)}}. \qquad \text{(A.35)}
\end{aligned}$$

Recall that:

$$\mathbb{V}\left(\mathcal{P}_{t+h|t+h} \,|\, \underline{\mathcal{F}_t}\right) = \left(\frac{1-\phi^{2h}}{1-\phi^2}\right)\left[\mathcal{K} + (1-\mathcal{K})\phi^2\right]\mathcal{K}\Sigma.$$

Thus, the ratio of the relative importance of factors is given by:

$$\frac{\mathbb{V}\left(\mathcal{P}_{2,t+h|t+h} \mid \underline{\mathcal{F}_t}\right)}{\mathbb{V}\left(\mathcal{P}_{1,t+h|t+h} \mid \underline{\mathcal{F}_t}\right)} = \left(\frac{1-\phi_2^{2h}}{1-\phi_1^{2h}}\right)\left(\frac{1-\phi_1^2}{1-\phi_2^2}\right)\frac{\left(\mathcal{K}_2+(1-\mathcal{K}_2)\phi_2^2\right)\mathcal{K}_2\Sigma_2}{\left(\mathcal{K}_1+(1-\mathcal{K}_1)\phi_1^2\right)\mathcal{K}_1\Sigma_1}\,. \tag{A.36}$$

The effect of h The only source of variation with respect to h is given by the first ratio, $\frac{1-\phi_2^{2h}}{1-\phi_1^{2h}}$. Let $v_h = \frac{1-\phi_2^{2h}}{1-\phi_1^{2h}}$, where $|\phi_1| < 1$ and $|\phi_2| < 1$. We look for the variation of v_h with respect to h. Let us denote by $a = \phi_2^2$ and $b = \phi_1^2$. We compute:

$$v_{h+1} - v_h = \frac{\left(1-a^{h+1}\right)\left(1-b^h\right) - \left(1-b^{h+1}\right)\left(1-a^h\right)}{\left(1-b^{h+1}\right)\left(1-a^h\right)}\,.$$

This ratio has the same sign as the numerator. We then note that:

$$1 - a^{h+1} = (1-a^h)a + 1 - a\,,$$

and similarly for b. Thus, the numerator transforms into:

$$\begin{aligned} v_{h+1} - v_h &\propto \left(1-b^h\right)\left[a\left(1-a^h\right)+1-a\right] - \left(1-a^h\right)\left[b\left(1-b^h\right)+1-b\right] \\ &\propto a + \frac{1-a}{1-a^h} - \left(b + \frac{1-b}{1-b^h}\right)\,. \end{aligned} \tag{A.37}$$

Let us denote by $f_h(a) = a + \frac{1-a}{1-a^h}$. Then $v_{h+1} - v_h$ has the same sign as the differential $f_h'(a)$ when a and b are close. This differential is given by:

$$f_h'(a) = 1 + \frac{ha^{h-1}(1-a) - (1-a^h)}{(1-a^h)^2}\,. \tag{A.38}$$

Notice that since $|a| < 1$, $(1-a^h)^2 < 1$, we can write the following inequality:

$$f_h'(a) > 1 + ha^{h-1}(1-a) - (1-a^h) = ha^{h-1}(1-a) + a^h > 0, \tag{A.39}$$

Thus the function is increasing in a for any order h. Eventually, we obtain that Δv_{h+1} is positive, so v_h is increasing whenever $a > b$, that is when $\phi_2 > \phi_1$, and vice versa.

We have shown above that the ratio $(1-\phi_2^{2h})/(1-\phi_1^{2h})$ increases (decreases) with maturity whenever $\phi_2 > \phi_1$ ($\phi_2 < \phi_1$). Thus, given a fixed maturity m, the weight $\omega_h^{(m)}$ varies inversely with $(1-\phi_2^{2h})/(1-\phi_1^{2h})$, that is $\omega_h^{(m)}$ grows with h when the first factor is more persistent than the second factor, i.e., $\phi_1 > \phi_2$.

The effect of maturity m We now look at the effect of the maturity on the variance decomposition. According to Equation (A.35), the relationship of the variance decomposition with respect to maturity is driven by $(B_{2,m}/B_{1,m})^2$. We look for a closed-form expression of these loadings.

Provided our specification of the market prices of risk (Equation 16), the risk-neutral dynamics of $\mathcal{P}_{t|t}$ are Gaussian and the autoregressive matrix is given by:

$$\Phi^{\mathbb{Q}} = \begin{pmatrix} \phi_1^{\mathbb{Q}} & \phi_{1,2}^{\mathbb{Q}} \\ 0 & \phi_2^{\mathbb{Q}} \end{pmatrix}.$$

Since the model is linear and conditionally Gaussian, it is easy to show that the price of a bond is given by:

$$\exp\left(-m \times y_t^{(m)}\right) = \exp\left(\mathcal{A}_m + \mathcal{B}_m' \mathcal{P}_{t|t}\right),$$

where

$$\mathcal{B}_m = \Phi^{\mathbb{Q}'} \mathcal{B}_{m-1} - \begin{pmatrix} \beta \\ 0 \end{pmatrix} = -\left(I - \Phi^{\mathbb{Q}'}\right)^{-1} \left(I - \Phi^{\mathbb{Q}'^m}\right) \begin{pmatrix} \beta \\ 0 \end{pmatrix}. \tag{A.40}$$

Tedious algebraic manipulations show that:

$$\mathcal{B}_m = \begin{pmatrix} -\frac{1-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}}}\beta & \frac{\phi_{1,2}^{\mathbb{Q}}}{1-\phi_2^{\mathbb{Q}}}\left[\frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{\phi_2^{\mathbb{Q}}-\phi_1^{\mathbb{Q}}} - \frac{1-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}}}\right]\beta \end{pmatrix}'. \tag{A.41}$$

Remember that $B_m = -\frac{1}{m}\mathcal{B}_m$, such that $\frac{B_{2,m}}{B_{1,m}} = \frac{\mathcal{B}_{2,m}}{\mathcal{B}_{1,m}}$. We therefore compute:

$$\frac{\mathcal{B}_{2,m}}{\mathcal{B}_{1,m}} = \frac{\phi_{1,2}^{\mathbb{Q}}}{1-\phi_2^{\mathbb{Q}}}\left[1 - \frac{1-\phi_1^{\mathbb{Q}}}{\phi_2^{\mathbb{Q}}-\phi_1^{\mathbb{Q}}} \times \frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}}\right]. \tag{A.42}$$

Let us first focus on $\frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}}$. It is useful to express it as:

$$\frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} = \frac{\phi_2^{\mathbb{Q}^m}-1+1-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} = 1 - \frac{1-\phi_2^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}}. \tag{A.43}$$

Using the result of Section A.5.4, we know that $(1-\phi_2^{\mathbb{Q}^m})/(1-\phi_1^{\mathbb{Q}^m})$ grows with m if $\phi_2^{\mathbb{Q}} > \phi_1^{\mathbb{Q}}$, and vice versa. Let us assume $\phi_2^{\mathbb{Q}} > \phi_1^{\mathbb{Q}}$. Then we have that:

$$\frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} = 1 - \frac{1-\phi_2^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} \quad \text{decreases with } m \text{ , } and$$

$$1 - \frac{1-\phi_1^{\mathbb{Q}}}{\phi_2^{\mathbb{Q}}-\phi_1^{\mathbb{Q}}} \times \frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} \quad \text{increases with } m\,.$$

Thus, $\left(\frac{\mathcal{B}_{2,m}}{\mathcal{B}_{1,m}}\right)^2$ increases with m, whatever the sign of $\varphi_{1,2}^{\mathbb{Q}}$ since it is squared. Let us now consider $\phi_2^{\mathbb{Q}} < \phi_1^{\mathbb{Q}}$. Then we have that:

$$\frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} = 1 - \frac{1-\phi_2^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} \quad \text{increases with } m\,,$$

and

$$1 - \frac{1-\phi_1^{\mathbb{Q}}}{\phi_2^{\mathbb{Q}}-\phi_1^{\mathbb{Q}}} \times \frac{\phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m}}{1-\phi_1^{\mathbb{Q}^m}} \quad \text{increases with } m \text{ since } \phi_2^{\mathbb{Q}^m}-\phi_1^{\mathbb{Q}^m} < 0\,.$$

Thus, $(\mathcal{B}_{2,m})/(\mathcal{B}_{1,m})^2$ increases with m as well. This shows that whatever the relative risk-neutral persistence of the processes, as long as $\phi_{1,2}^{\mathbb{Q}} \neq 0$, the ratio $(\mathcal{B}_{2,m})/(\mathcal{B}_{1,m})^2$ grows with the maturity m. The limiting values are:

$$\left(\frac{\mathcal{B}_{2,1}}{\mathcal{B}_{1,1}}\right)^2 = 0 \quad \text{and} \quad \left(\frac{\mathcal{B}_{2,\infty}}{\mathcal{B}_{1,\infty}}\right)^2 = \frac{\phi_{1,2}^{\mathbb{Q}^2}}{\left(1-\phi_2^{\mathbb{Q}}\right)^2}. \tag{A.44}$$

In the end, since $(\mathcal{B}_{2,m})/(\mathcal{B}_{1,m})^2$ is always increasing, we have that $\omega_h^{(m)}$ **decreases** with the maturity (see Equation A.35). Thus, since the variance decomposition of the second factor $\mathcal{V}_2$ is higher than that of the first one, the relative importance of macro news grows with maturity. This is consistent with what we obtain empirically.

A.6 Aggregation Schemes

Put in general terms, our goal is to construct the monthly time series of $\Delta\mathcal{P}_t$ based on the high-frequency measures $\Delta F_{t,\tau,j}$ within each month. All variations of our method consist in successive linear regressions.

A.6.1 Including RV in the regression Define the monthly 3×3 realized variance-covariance matrix RV_t computed from daily data. Then we include the unique elements of RV_t in a regression extending Equation (A.17) as follows:

$$\Delta\mathcal{P}_t^{(i)} = \alpha_i + \gamma_i \cdot \mathcal{J}_t + \beta_i^\top \left[\sum_{j=1}^{J} \sum_{\tau\in[t,t+1)} \Delta F_{t,\tau,j}\right] + \Gamma_i^\top \, Vech\left(\mathrm{RV}_t\right) + \eta_t^{(i)}\,, \tag{A.45}$$

where $Vech$ is the half vectorization operator. We estimate the specification (A.45) by OLS and consider the fitted values to be the news component.

A.6.2 Correcting for the efficiency loss with WLS Assume that the following relationship holds:

$$\Delta\mathcal{P}_t = \alpha + \Lambda \sum_{\tau\in[t,t+1)} \Delta\mathcal{P}_\tau, \tag{A.46}$$

but that the high-frequency impact is measured with errors: $\Delta\mathcal{P}_\tau = \widehat{\Delta\mathcal{P}_\tau} + \eta_\tau$. Then, we have that:

$$\Delta\mathcal{P}_t = \alpha + \Lambda \sum_{\tau\in[t,t+1)} \widehat{\Delta\mathcal{P}_\tau} + \Lambda \sum_{\tau\in[t,t+1)} \eta_\tau. \tag{A.47}$$

Therefore, the residuals of the monthly regressions $\nu_t = \Lambda \sum_{\tau\in[t,t+1)} \eta_\tau$ inherit the potential heteroskedasticity of η_τ. One way to control for this heteroskedasticity is to perform weighted least squares in the second stage, assuming that $\nu_t = \Lambda\Omega_t^{1/2}\widetilde{\nu_t}$. In practice, we measure Ω_t by the following realized covariance estimate:

$$\Omega_t = \frac{1}{d_t - 1} \sum_{\tau\in[t,t+1)} \widehat{\eta_\tau}\widehat{\eta_\tau}^\top, \tag{A.48}$$

and $\Omega_t^{1/2}$ is the Cholesky decomposition of Ω_t. Notice that the problem of the covariance of the residuals has not been entirely treated, since Λ is unknown before running the regression. To address the this, we use the following estimation procedure. Denote by $\Delta\mathcal{P}$ the vector of size $3T$ stacking all time observations of $\Delta\mathcal{P}_t$, define $\mathbf{X}$ as follows:

$$\mathbf{X} = \left[1, \sum_{\tau\in[t,t+1)} \widehat{\Delta\mathcal{P}_\tau}\right] \otimes I_3, \tag{A.49}$$

and write the second-stage regression as follows:

$$\Delta\mathcal{P} = \Gamma + \nu, \tag{A.50}$$

where $\Gamma = [\alpha^\top, Vec(\Lambda)^\top]^\top$ and the variance-covariance matrix of the residuals is a rotation of:

$$\Omega = bdiag\left(\Omega_1, \ldots, \Omega_T\right), \tag{A.51}$$

where $bdiag$ is the block-diagonal operator, building a large block-diagonal matrix out of the (3×3) realized covariance matrices Ω_t. Defining $\Omega = C_\Omega C_\Omega^\top$ the block-diagonal Cholesky decomposition of Ω, our estimator of Γ is given running the OLS regression:

$$\begin{aligned} & C_\Omega^{-1}\Delta\mathcal{P} = C_\Omega^{-1}\mathbf{X}\,\Gamma + \nu_t^* \\ \iff\quad & \Delta\mathcal{P}^* = \mathbf{X}^*\,\Gamma + \nu_t^*. \end{aligned} \tag{A.52}$$

Note that ν_t^* still has variance covariance given by $\Lambda\Lambda^\top$ in principle. We measure its variance covariance by assuming it is constant over time, and with the estimator:

$$\widehat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \widehat{\nu_t^*}\,\widehat{\nu_t^*}^\top . \tag{A.53}$$

We then perform a second stage regression in the same fashion, denoting by C_Σ the Cholesky decomposition of $\widehat{\Sigma}$:

$$\begin{aligned} & C_\Sigma^{-1}\Delta\mathcal{P}^* = C_\Sigma^{-1}\mathbf{X}^*\,\Gamma + \nu_t^{**} \\ \iff \quad & \Delta\mathcal{P}^{**} = \mathbf{X}^{**}\,\Gamma + \nu_t^{**} . \end{aligned} \tag{A.54}$$

Last, we form the series of news components of the monthly PCs simply by computing $\mathbf{X}\widehat{\Gamma}$, where $\widehat{\Gamma}$ is the last estimated set of parameters from Equation (A.54).

A.6.3 Disaggregating news types Instead of directly transforming the futures data, we can consider daily regressions as a first step, and the aggregation as a second step. Consider the following regression of the *daily* principal components:

$$\Delta\mathcal{P}_{t,\tau}^{(i)} = \sum_{j=1}^{J} \beta_{i,j}^\top \left[\Delta F_{t,\tau,j} - \mathbb{E}_{\tau-1}\left(\Delta F_{t,\tau,j}\right)\right] + \eta_{t,\tau}^{(i)} . \tag{A.55}$$

We assume that each news type has its own premium, which is invariant through time, such that:

$$a_{t,\tau,j-1}^{(\mu)} = a_j^{(\mu)} . \tag{A.56}$$

The specification then transforms:

$$\Delta\mathcal{P}_{t,\tau}^{(i)} = \sum_{j=1}^{J} \beta_{i,j}^\top \left[\Delta F_{t,\tau,j} - a_j \cdot \mathbf{1}\{\text{news } j \text{ at } \tau\}\right] + \eta_{t,\tau}^{(i)} . \tag{A.57}$$

The estimates can be obtained by regressing the variation of the daily PCs onto all $3J$ daily series of futures and the J series of dummy variables indicating whether there is a news of type j on day τ. We estimate this specification by OLS and aggregate the results at the monthly frequency by performing the following regression:

$$\begin{aligned} \Delta\mathcal{P}_t^{(i)} &= \alpha_i + \lambda_i^\top \sum_{\tau\in[t,t+1)} \widehat{\Delta\mathcal{P}_{t,\tau}} + \nu_t^{(i)} \\ &= \alpha_i + \sum_{\ell=1}^{3} \lambda_{i,\ell} \left[\sum_{\tau\in[t,t+1)} \sum_{j=1}^{J} \widehat{\beta_{\ell,j}}^\top \left[\Delta F_{t,\tau,j} - \widehat{a_j}\cdot\mathbf{1}\{\text{news } j \text{ at } \tau\}\right]\right] + \nu_t^{(i)}. \end{aligned} \tag{A.58}$$

This second step is once again estimated by OLS and the fitted values of this regression define the news component of the monthly PCs.

A.6.4 Disaggregating news types and sign This specification combines the monthly separation of news impact depending on their signs and the daily specification allowing for news-type disaggregation. Our specification of Equation (A.57) is replaced by:

$$\Delta\mathcal{P}_\tau^{(i)} = \sum_{j=1}^{J} \beta_{i,j,+}^\top \left[\Delta F_{t,\tau,j}^+ - a_j^+ \cdot \mathbf{1}\{\text{news } j \text{ at } \tau\}\right] + \beta_{i,j,-}^\top \left[\Delta F_{t,\tau,j}^- - a_j^- \cdot \mathbf{1}\{\text{news } j \text{ at } \tau\}\right] + \eta_\tau^{(i)}. \tag{A.59}$$

With our assumptions, only the sum $a_k^+ + a_k^-$ is identified, and the estimates can be obtained by running OLS of the variation of the daily PCs onto all $3J$ daily series of positive futures changes, the $3J$ daily series of negative futures changes and the J series of dummy variables indicating whether there is a news of type j on day τ. We aggregate the results at the monthly frequency in the exact same fashion similar to Equation (A.58):

$$\Delta\mathcal{P}_t^{(i)} \quad = \quad \alpha_i + \lambda_i^\top \sum_{\tau\in[t,t+1)} \widehat{\Delta\mathcal{P}_\tau} + \nu_t^{(i)}. \tag{A.60}$$