DISCUSSION PAPER SERIES

DP17930

THE CENTRAL BANK CRYSTAL BALL: TEMPORAL INFORMATION IN MONETARY POLICY COMMUNICATION

David Byrne, Robert Goodhead, Michael McMahon and Conor Parle

MONETARY ECONOMICS AND FLUCTUATIONS

CEPR

ISSN 0265-8003

THE CENTRAL BANK CRYSTAL BALL: TEMPORAL INFORMATION IN MONETARY POLICY COMMUNICATION

David Byrne, Robert Goodhead, Michael McMahon and Conor Parle

Discussion Paper DP17930
Published 21 February 2023
Submitted 21 February 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

THE CENTRAL BANK CRYSTAL BALL: TEMPORAL INFORMATION IN MONETARY POLICY COMMUNICATION

Abstract

Effective central bank communication provides information that the public wants but does not have. Using a new textual methodology to quantify the temporal information in central bank communication, we argue that central bank assessments of the (latent) state of the economy can be the source of the public's information deficit, rather than superior information necessarily. The implication of this is that communication of a single, fixed, reaction function, even if desirable, is likely impossible even if preferences remain fixed over time. Communication of how the central bank is assessing the economy should be emphasised in addition to any forward guidance.

JEL Classification: E52, E58, C55

Keywords: Communication

David Byrne - david.byrne@centralbank.ie
Central Bank Of Ireland

Robert Goodhead - robert.goodhead@centralbank.ie
Central Bank Of Ireland and CEPR

Michael McMahon - michael.mcmahon@economics.ox.ac.uk
Department of Economics, University of Oxford and CEPR

Conor Parle - Conor.Parle@ecb.europa.eu
European Central Bank

Acknowledgements
McMahon gratefully acknowledges financial support from the European Research Council (Consolidator Grant Agreement 819131). We thank Mengqi Qiu, Delia Macaluso and Tim Williams for excellent research assistance. We are grateful to Ana Aguilar, Michael Bauer, Michael Ehrmann, Sharon Kozicki, Reamonn Lydon, Athanasios Orphanides, Davide Romelli, and Ulf Soderstrom for helpful comments and discussions, as well as seminar participants at numerous workshops, seminars and conference for helpful comments. All the algorithms used in this paper are available from the authors' webpages and github. The views presented in the paper are those of the authors alone and do not represent the official views of the Central Bank of Ireland or the European System of Central Banks. Any errors remain ours alone.

The Central Bank Crystal Ball: Temporal information in monetary policy communication*

David Byrne[†], Robert Goodhead[‡], Michael McMahon[§] Conor Parle[¶]

First Draft: May 6, 2021
This Draft: February 21, 2023

Abstract

Effective central bank communication provides information that the public wants but does not have. Using a new textual methodology to quantify the temporal information in central bank communication, we argue that central bank assessments of the (latent) state of the economy can be the source of the public's information deficit, rather than superior information necessarily. The implication of this is that communication of a single, fixed, reaction function, even if desirable, is likely impossible even if preferences remain fixed over time. Communication of how the central bank is assessing the economy should be emphasised in addition to any forward guidance.

Keywords: Monetary Policy, Communication, Natural Language Processing

JEL Codes: E52, E58, C55

*McMahon gratefully acknowledges financial support from the European Research Council (Consolidator Grant Agreement 819131). We thank Mengqi Qiu, Delia Macaluso and Tim Williams for excellent research assistance. We are grateful to Ana Aguilar, Michael Bauer, Michael Ehrmann, Sharon Kozicki, Reamonn Lydon, Athanasios Orphanides, Davide Romelli, and Ulf Söderström for helpful comments and discussions, as well as seminar participants at numerous workshops, seminars and conference for helpful comments. All the algorithms used in this paper are available from the authors' webpages and github. The views presented in the paper are those of the authors alone and do not represent the official views of the Central Bank of Ireland or the European System of Central Banks. Any errors remain ours alone.

[†]Central Bank of Ireland. Email: david.byrne@centralbank.ie

[‡]Central Bank of Ireland. Email: robert.goodhead@centralbank.ie

[§]University of Oxford, CEPR, CfM (LSE). Email: michael.mcmahon@economics.ox.ac.uk

[¶]European Central Bank and Trinity College Dublin. Email: Conor.Parle@ecb.europa.eu

1 Introduction

Does the fact that asset prices move in response to central bank statements mean markets are learning from the central bank crystal ball? A burgeoning literature uses high-frequency asset price movements around communication events to provide evidence that monetary policymakers influences market expectations when they communicate (for instance, studies such as Kuttner 2001, Gürkaynak et al. 2005 and Gertler and Karadi 2015). Such studies provide strong evidence of market participants having an "*information deficit*" relative to policymakers, in the sense that market participants are unable to fully anticipate what a central banker will say, prior to their statements. This information deficit is, at least partly, filled by central bank statements from which useful information for the formation of expectations is garnered. But what is the source of the information about the future that the central bank provides such that markets can be influenced by its pronouncements? Understanding the answer to this is vital in order to be able to think about what and how a central bank should optimally communicate.

In many models used for the analysis of monetary policy, the assumptions of full information and rational expectations (FIRE), such as in Eggertsson and Woodford (2003) or Galí (2008), mean that the central bank only usefully communicates two things. First, it communicates its reaction function, although when the central bank is assumed to commit, they only have to do this once. Second, it can communicate forward guidance in the form of planned deviations from the reaction function. In models of asymmetric information, it can also communicate its superior economic information.

In this paper, we highlight the role of conjunctural economic assessment in how communication affects asset prices. This assessment role does not arise in standard models though it is a core idea in the work on real-time monetary policymaking (Orphanides 2001). It captures the mapping from observed macroeconomic and financial data to the underlying (latent) state of the economy such as the shocks that have hit the economy, their persistence and the state of the output gap. This mapping is a key source of uncertainty facing the central bank, markets and other economic agents, even where they are only concerned with the forward-looking nature of monetary policy.

The potentially important role of private knowledge about the economy for yield curve movements is evident in many papers considering the information effect as in, for example, Romer and Romer (2000), Ellingsen and Söderström (2001), and Nakamura and Steinsson (2017). Other papers show that the information could be related to communication of uncertainty (Hansen et al. 2019, Cieslak and Schrimpf 2019). Our focus on conjunctural assessment is consistent with these approaches though, importantly, we argue that communicating a different assessment does not necessarily reflect *superior* information. Instead, it constitutes an important element of communicating the reaction function. And overlooking its role, as the literature has typically done, risks unduly focusing on the role of forward guidance.

Our main contribution is empirical. We study new measures of both the communication transmission mechanism, and the information deficit. It is part of a complementary empirical approach that begins with the content of statements and relates these to the asset price movements. By directly quantifying the extent of any information deficit between market participants and policymakers, this approach can foster a greater understanding of the communication transmission mechanism.

Specifically, in this paper we contribute to the literature in four main ways. First, we present a framework to describe monetary policy decision making, and the market anticipation of the decisions, that is more general than the one that arises from standard monetary DSGE models. *A priori*, information regarding the future might be expected to be especially important for financial markets, relative to information about the past. This is especially true if we believe that policymakers and markets have similar access to the necessary information on the state of the economy. However, our description of the monetary policy decision-making process shows that, even with the same basic information, central banks assessment of economic conditions and their likely evolution are both potentially important sources of the information deficit. This conjunctural assessment relies on both an *evaluation* of the current state of the economy based on past data, and a *projection* from the current state to the future.

Second, we develop an algorithm to quantify the temporal dimensions of the central bank communication. Our methodology adapts new tools from the Natural Language Processing (NLP) literature in order to suit our central bank communication context. We validate our approach using labelled Federal Reserve text from separate parts of the Tealbook briefings; our methodology maps well to the framework we outline. This algorithm allows us to understand the role and importance of temporal information in the the communication transmission mechanism. While there has been a growing emphasis on the importance of communicating the future intentions of monetary policy makers as part of the management of expectations, studies into the explicit importance of temporal information in textual data are, surprisingly, rare to date. This algorithm, potentially following topic-specific adjustments, is likely to be useful in other applications. In an accompanying paper, Byrne et al. (2023), we explore the details of the algorithm in a broader setting and make clear where other researchers may wish to adapt, extend, or revise our approach to fit different contexts.

Having successfully tagged cases of temporal information in our corpora, our third contribution is to show that the importance this information for both the Federal Reserve and the ECB. Perhaps unsurprisingly, we find that information regarding the future has an important role at explaining asset price movements, and we interpret this as evidence in favour of the idea that markets react to information in central bank forecasts and their forward guidance. More surprisingly, we show that even backward-looking data is extremely informative. This, we argue, captures important contextualisation of the data contained in the central bank's economic assessment to which markets react.

Our fourth contribution is to develop a novel measure of the information deficit itself using the questions raised by journalists after monetary policy announcements. This complements the NLP approach with more direct evidence showing that the information deficit includes a significant conjunctural assessment dimension. For this analysis, we focus on the ECB which has the advantage of having a longer, and more consistent, approach to post-meeting press conferences than the Fed. The idea is that, having heard the opening statement, financial journalists query policymakers to fill in gaps that their readers will have. These are, to the best of our knowledge, the first direct measures of the information deficit.

To validate this interpretation, we show that the more closely a speech is related to the questions asked after the most-recent statement, the stronger is the market reaction indicating greater news and belief updating. This suggests that questions contain important signals of the information that markets wish to hear about; that is, they reflect the information deficit between the ECB and financial markets. We then show that these questions are not just about the ECB's forward guidance, but, as suggested by our framework and NLP analysis, also concern the conjunctural assessment.

Why does the finding that markets garner important information from the central bank's assessment of recently released (mostly common) information matter? This result is important because it identifies one of the key sources of uncertainty that central banks face. Orphanides (2001) highlighted the difficult task that central bank faced making policy in real time by showing that policy reaction functions estimated using real-time data differ considerably from those estimated with ex-post, revised data. But our paper also goes further in showing that the filtering of real-time data into signals about the state of the economy seems to be as much or more of a problem for markets, and that markets react to how the central bank is viewing the world.

Our findings also matter for discussions of rules versus discretion, and the role of communication. Kocherlakota (2016) shows that discretion can be optimal in environments where there is a very large set of economic variables that matter for assessing the state of the economy and optimal policy, making it impossible to contract on the full set of possible variables. In FIRE models, where there is knowledge of the full structure of the economy, central bank credibility is assured by a commitment to a sufficiently aggressive linear inflation rule. By contrast, Orphanides and Williams (2005) showed that the job of anchoring inflation expectations is more difficult when agents do not understand the structure of the economy and/or the central bank's rule, or if the central bank has some superior information. This is because the link between economic outcomes and expectations is broken. In such an environment, agents may react to the assessments of the state of the economy, or their outlook for the future, and the central bank may have to regularly renew its vow to aim to achieve its objective of price stability. Communication takes on a much broader role.

The rest of the paper is structured as follows. In section 2, we present our framework for how monetary policy is set and then we relate this framework to the temporal dimension of the policy decision. We then outline our NLP methodology to measure the temporal dimension of policy communications (section 3). Our NLP analysis in section 4 focuses on both the Federal Reserve and the ECB. In the penultimate section of the paper (section 5), we analyse develop our novel measure of the information deficit using journalists questions at the ECB's post-meeting press conferences. Section 6 concludes.

2 Monetary Policy Surprises and Information Deficits

In this section we show that the complexity of the monetary policy process can be grouped into three broad aspects. Each is a potential source of news for market participants. Although monetary policy is inherently forward-looking, we show that the source of news that generates market surprises could as reasonably come from backward-looking assessments of the current data context as from the forward-looking forecast.

2.1 Monetary Policymaking Process

Consider a monetary policymaker in month m who has to decide the interest rate, i_m. They have access to a large amount of data capturing macroeconomic trends, surveys, market prices, and other relevant information. We represent this as a high dimensional vector X_m^{CB}. Their decision-making process can be summarised in two broad steps:

1. The policymaker maps the data into a vector of beliefs about the current state of the economy: $\Omega_m^{CB} = g_m\left(X_m^{CB}\right)$ where $g_m(\cdot)$ captures the analysis of the data available at meeting m, as well as the forecast (including the judgement applied).

2. The policymaker selects the appropriate interest rate as a function of this state to reflect their preferences: $i_m = f_m\left(g_m\left(X_m^{CB}\right)\right)$

The reaction function maps the objective economic data to the desired interest rate. In our case it is given by the nested function: $f_m\left(g_m\left(X_m^{CB}\right)\right)$. In many standard reaction functions, the states to which monetary policy reacts consist of inflation deviations from target and the output gap ($X_m^{CB} = [\hat{\pi}_m, \hat{y}_m]$). This is because in FIRE models these contemporaneous state variables are assumed to be perfectly observable and, since the structural parameters of the economy are known with certainty, the mapping from the current state of the economy to the optimal forecast is straightforward. The current state variables are sufficient statistics for the optimal h-period ahead forecasts ($\hat{\pi}^e_{m,m+h}, \hat{y}^e_{m,m+h}$), leaving no role for assessment of the state through $g_m(\cdot)$.

In FIRE models, $f_m(\cdot)$ is typically linear, reflecting the linearity of the underlying model, and does not vary with time, reflecting the assumption of central bank commitment. Moreover, an exogenous monetary policy shock is often added linearly to the endogenous reaction terms; in our framework, $f_m\left(g_m\left(X_m^{CB}\right)\right) - f_{m-}\left(g_m\left(X_m^{CB}\right)\right)$ would capture the same kind of shift in policy for a given state of the economy. The difference is that, in our case, the cause of the shift may be endogenous to economic conditions, as in McMahon and Munday (2022a).

In our generic description, we have allowed both elements of the reaction function to vary each meeting. Carney (2017) argues that time-varying reaction functions are completely natural in the practice of monetary policy, even if they are not standard in our models.[1] Clarida et al. (1999) and Sims and Zha (2006), among many others, find evidence for low-frequency changes in the reaction function in the US. McMahon and Munday (2022b) present some updated evidence for time-varying reaction functions in the US, while the role played by reaction function variation for market surprises is emphasised by McMahon and Munday (2022a). While we acknowledge that possibility in this paper, and we may capture some of this variation with our textual measures, reaction function variation is not the main focus of this paper.

We focus on the aspect of the policy process that has received less attention in macro models: $g_m(\cdot)$. As should be clear to central bankers, this function captures the very heart of the interest rate process, and the work of economists in central banks. And it is not just of practical relevance. Even if the central bank, as is desirable in many models of optimal policy, could commit to its preferences ($f_m(\cdot) = f_{m-1}(\cdot)\ \forall m$), and communicate those perfectly, the presence of a time-varying $g_m(\cdot)$ assessment function means that it would be impossible to have a time-invariant mapping between macro data X_m and interest rate choices. This means that identified time-variation in estimated reaction functions *could* reflect evolving assessments of the economy, rather than changing preferences.[2]

[1] The cause of such changing reaction functions includes changing membership of the monetary policy committee, central bank uncertainty about the structure of the economy which changes over time, and time variation in the persistence of shocks.

[2] $g_m(\cdot)$ is more than just a mapping from contemporaneous data to the forecasts for those data. The mapping between, for example, forecasts for inflation and other macro variables may vary depending on assessments about the type and persistence of the shocks that have hit the economy.

The assessment function captures a different type of uncertainty than the typical linearly separable shock that we typically include in models. Those shocks represent *aleatoric* uncertainty, relating to the outcomes of random processes. Consistent with the growing literature on narrative models (Shiller 2017, Larsen et al. 2021, Goetzmann et al. 2022, Andre et al. 2022, Flynn and Sastry 2022), we can think of the assessment function as reflecting *ontological* uncertainty[3]. That is, without perfect knowledge of the true data-generating model, different agents process the same generated data through the models in their own heads, and can reach different conceptualizations of the current state of the economy and its likely evolution. Macaulay and Song (2022) find that multiple distinct economic narratives often circulate about the same economic events. Such ontological uncertainty is exacerbated by *epistemic* uncertainty: the occurrence of events or shocks that have not been seen before enough, or ever, to understand their effects properly.

In summary, the assessment function $g_m(\cdot)$ acknowledges and represents the more general forms of uncertainty faced by policymakers, and those who try to interpret policymakers' communications, than can be represented in models that exclude it.

2.2 Information Deficits and Market Surprises

As it is an important novelty in this paper, we will return to discuss the assessment function in more detail below. But first we show how understanding the policy process can help us to understand the sources of market surprises.

We assume that market participants are interested in what policies the central bank will adopt in the future as in Ellingsen and Söderström (2001). That is, they form expectations of the central bank policy choice and these are reflected in market prices of various securities. We denote the average expectation just before the announcement of the monetary policy decision at time m as $\mathbb{E}\left[\, i_t \;\middle|\; \mathcal{I}^{mkt}_{m-} \,\right] = \tilde{f}_{m-}\left(\tilde{g}_{m-}\left(X^{mkt}_{m-}\right)\right)$ where $m-$ indicates the moment just before the month m decision, $\mathcal{I}^{mkt}_{m-}$ is the information set at that moment, X^{mkt} is the economic data that the market has, and $\tilde{f}_m(.)$ and $\tilde{g}_m(.)$ are the market beliefs about the central bank equivalent functions described earlier.

A monetary event, such as an announcement of policy and the associated communications, provides information from the central bank. The new, updated information is $\mathcal{I}^{mkt}_{m}$ and the expectation is $\tilde{f}_m\left(\tilde{g}_m\left(X^{mkt}_m\right)\right)$. The market surprise, as measured in the empirical literature is, therefore, $\varepsilon_m = \mathbb{E}\left[\, i_m \;\middle|\; \mathcal{I}^{mkt}_{m} \,\right] - \mathbb{E}\left[\, i_{m-} \;\middle|\; \mathcal{I}^{mkt}_{m-} \,\right]$. This surprise could be generated by changes in any of the three broad factors in the central bank decision:

1. X^{CB}_m could be revealed to include new information not in X^{mkt}_m.

[3] A familiar form of ontological uncertainty for economists is identification uncertainty. Multiple model specifications can be put forward to capture an economic quantity of interest given the true data-generating process is unknown.

2. The central bank could provide an updated assessment/forecast of the economy. This would reveal if it has re-evaluated how it is assessing the state of economy and forecasting its evolution as captured by $g_m(\cdot) \neq \tilde{g}_{m-}(\cdot)$

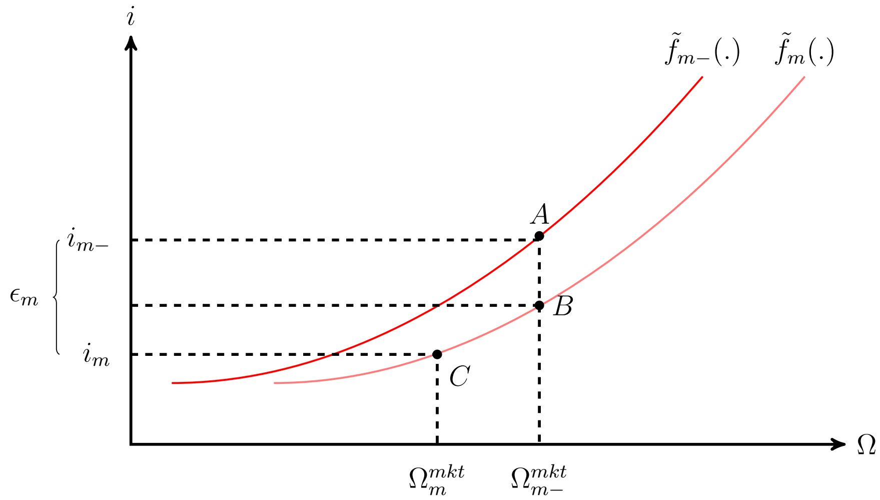
3. The central bank may choose to react to a given state of the economy more or less aggressively than previously; $f_m(\cdot) \neq \tilde{f}_{m-}(\cdot)$.

The market surprise is therefore given by: $\varepsilon_m \equiv \tilde{f}_m(\Omega_m^{mkt}) - \tilde{f}_{m-}(\Omega_{m-}^{mkt})$. To understand its drivers, we can approximately decompose it using a first-order Taylor series expansion of $\tilde{f}_m(\Omega_m^{mkt})$ around the pre-announcement view of the economic state Ω_{m-}^{mkt}:[4]

$$\varepsilon_m \approx \underbrace{\tilde{f}_m(\Omega_{m-}^{mkt}) - \tilde{f}_{m-}(\Omega_{m-}^{mkt})}_{\text{Updated Preferences}} + \underbrace{(\Omega_m^{mkt} - \Omega_{m-}^{mkt})\tilde{f}'_m(\Omega_{m-}^{mkt})}_{\text{Reassessment of State}} \tag{1}$$

To see this decomposition visually, consider Figure 1. Point A is the pre-announcement expectation for the interest rate ($i_{m-} = \tilde{f}_{m-}(\Omega_{m-}^{mkt})$); Point C captures the expectations post-announcement ($i_m = \tilde{f}_m(\Omega_m^{mkt})$).[5] The effect of the updated preferences is captured by the move from A to B, while B to C reflects the impact of the reassessment of the state of the economy through the updated preferences.[6]

Figure 1: Market Surprise Decomposition



[4] $\tilde{f}_m(\Omega_{m-}^{mkt}) \approx \tilde{f}_m(\Omega_m^{mkt}) + (\Omega_m^{mkt} - \Omega_{m-}^{mkt})\tilde{f}'_m(\Omega_{m-}^{mkt})$. Subtracting this from $\tilde{f}_m(\Omega_m^{mkt})$ yields the expression in the text.

[5] Following the interest rate announcement, there is knowledge of the revealed policy interest rate but markets still have to form expectations for longer maturity rates.

[6] Of course, where the preferences are non-linear, as depicted, other decompositions between and A and C are possible and each will give slightly different weight to each component.

2.3 The Temporal Dimension of Economic Assessment

The economic assessment function $g_m(\cdot)$ is not widely studied in the context of empirical and theoretical studies of monetary policy. When one considers the degree of uncertainty about the structure of the economy, which is constantly evolving, the crucial importance of assessing the state of the economy is clear. Even the data on the state of the economy is noisy and released with substantial lags. As such, central banks are constantly monitoring data, adjusting their views of what shocks have hit the economy and modifying their forecast judgements as part of the policy process. This is why central banks invest heavily in the process of nowcasting and in attempts to uncover the state of latent variables like the output gap.

As discussed above, a FIRE environment admits little role for communication other than of a commitment to deviate policy from one's reaction function and forward guidance of deviations from this rule. Though the central bank looks into its crystal ball to forecast, central banks and markets in such environments form the same expectations. In reality, there are more dimensions than this through which a central bank can influence markets with its communication, and the assessment function is at the core of several of these.

We argue that assessment communications are inherently temporal in nature: including both forward-looking dimensions, such as projections, and backward-looking dimensions, such as interpretations of the data. To see this, we further split the assessment function, $g_m\left(X_m^{CB}\right)$ into two distinct steps:

Evaluation: This involves following developments across a range of domestic and international markets using macroeconomic, financial and other data. Recent data is analysed and put in context in order to interpret its movements (e.g., is a recent increase in inflation likely to be transitory?). This analysis is backward-looking in the sense that it uses the latest available past data pertaining to recent developments, and also more historical data for context and within statistical analyses.

Specifically, let the conjunctural economic analysis function map observed (past) data into a view of the current economic state (and the shocks that have hit it):

$$\Sigma_m^{CB} = s_m(X_m^{CB}). \tag{2}$$

Projection: One reason for the existing focus on the forward guidance dimension of central bank communication is that monetary policy is inherently forward-looking. It is made on the basis of forecasts of key variables, and the risks around them. Forecasts typically use a suite of models, and judgement informed by the economic analysis, to map the current conjunctural assessment to the relevant future horizons:

$$\Omega_m^{CB} = p_m(\Sigma_m^{CB}). \tag{3}$$

Our description allows both analysis and forecasting functions to vary each meeting. $s_m(\cdot)$ might vary because unfolding data, including past data revisions, might change our view of what types of shocks have hit or how persistent they are likely to be. $p_m(\cdot)$ updates may reflect new forecast judgements about the likely future evolution of the economy. We will utilise Natural Language Processing techniques that measure temporality inherent in the Evaluation and Projection steps.

Since $g_m(\cdot) = p_m(s_m(\cdot))$, we can further decompose the market surprise:[7]

$$\begin{aligned}
\varepsilon_m \approx\ & \underbrace{\tilde{f}_m(\Omega_{m-}^{mkt}) - \tilde{f}_{m-}(\Omega_{m-}^{mkt})}_{\text{Updated Preferences}} \\
& + \underbrace{\left[\tilde{p}_m(s_{m-}(X_{m-}^{mkt})) - \tilde{p}_{m-}(s_{m-}(X_{m-}^{mkt}))\right] \tilde{f}'_m(\Omega_{m-}^{mkt})}_{\text{Reassessment: Projection}} \\
& + \underbrace{\left[\tilde{s}_m(X_{m-}^{mkt}) - \tilde{s}_{m-}(X_{m-}^{mkt})\right] \tilde{p}'_m(s_{m-}(X_{m-}^{mkt}))\tilde{f}'_m(\Omega_{m-}^{mkt})}_{\text{Reassessment: Evaluation}} \\
& + \underbrace{(X_m^{mkt} - X_{m-}^{mkt})\tilde{s}'_m(X_{m-}^{mkt})\tilde{p}'_m(s_{m-}(X_{m-}^{mkt}))\tilde{f}'_m(\Omega_{m-}^{mkt})}_{\text{Reassessment: New Information}}
\end{aligned} \tag{4}$$

As noted in the Introduction, an earlier literature discussed whether there is an important role for private information in monetary policy. It is true that central banks have *some* information that is not publicly available, such as regulatory data from financial firms, or sometimes have early access to data. Nonetheless, markets have access to almost all of the underlying data, and it would be rare for the central bank announcement to formally reveal whatever private information there is. If $X_m^{CB} \simeq X_{m-}^{mkt}$, genuinely new, private information (the third reassessment channel in Eqn. 3) would not be an important driver of the surprises (as argued by Bauer and Swanson 2020).

What is revealed, and what could be the source of perceived informational advantage, is a differential assessment of economic conditions. The central bank often has greater analytical resources in the form of large teams of economists and multiple models. This is captured by $s_m(\cdot)$ and reassessments of the economic state are often communicated. For example, this could be putting more or less weight on different sources of data at different times to better explain the observed developments. This is the backward-looking *evaluation* of the data. The central bank may also communicate its updated views and

[7] We sequentially applying a first-order Taylor series expansion on $\tilde{p}_m(.)$ and $\tilde{s}_m(.)$ around $s_{m-}(X_{m-}^{mkt})$ and X_{m-}^{mkt} respectively.

assumptions on the outlook for the economy over the forecast horizon. This is captured by $p_m(\cdot)$. For instance, an updated forecast might follow from having reassessed the nature and pattern of recent forecast errors. This is the future-focused *projection* dimension.

A typical example of the importance of assessment in communication comes from the monetary policy statement of the ECB in September 2019 (Draghi 2019). The text comprises 23 paragraphs, of which 17 relate to assessment ($g_m(\cdot)$) and the remainder to the mapping into the policy decision ($f_m(\cdot)$). Within the assessment portion of the text, all but two paragraphs relate to the ECB's re-assessed evaluation of the state ($s_m(\cdot)$). Those remaining two paragraphs cover the ECB's latest projections ($p_m(\cdot)$). Similarly for the Federal Reserve, a typical FOMC statement[8] begins with an evaluation of the recent data, followed by the policy decision, projections, and an expressed intention to base decisions at future meetings on re-evaluations of the state at those times.

These examples highlight the importance of the assessment function in central bank communication. More pertinently however, they emphasize the temporality embedded in the communication, and the particular weight given to backward-looking evaluation.

3 Measuring Temporal Communication

The previous section has established that market surprises may come from both the projection and the evaluation aspects of central bank decision making. We do not, at present, have a method to precisely identify communication corresponding to each stage of the assessment and policy preference functions outlined above; the high-dimensional nature of communication and the over-lapping nature of the language make this a difficult task. In order to examine the extent to which these channels matter in practice, we will conduct an empirical test of the importance of future and past temporal information.

Before we can examine the role of temporal information empirically, we have to measure it. While existing studies of central bank communication have emphasised two Ts, Topic and Tone, attention on the "third T", temporality, is limited. One reason for this is that extracting temporal information from text can be a demanding task. In this section, we present an algorithm that we have developed specifically for measuring the temporal dimension of central bank textual data. Before describing our algorithm, we will summarise briefly the existing economic literature focused on the quantification of textual data.[9]

[8]See June 2019 Federal Reserve Statement.

[9]A more detailed literature summary can be found in the companion paper, Byrne et al. (2023).

3.1 The Three Ts of Text Analysis

The NLP literature quantifying the tone of text has been widely adapted for economics. Early studies used manual coding of texts to measure tone (Ehrmann and Fratzscher 2009, Rosa and Verga 2007, Hayo and Neuenkirch 2010, Picault and Renault 2017), while others studies have used standard dictionaries (Schmeling and Wagner 2019, Cannon 2015). Some approaches have devised dictionaries specific to the language of central banking and monetary policy (Bennani and Neuenkirch 2017, Apel et al. 2019, Hubert and Labondance 2021, Hansen and McMahon 2016, Shapiro et al. 2020, Parle 2022) while others used non-dictionary methods (Lucca and Trebbi 2009, Tobback et al. 2017, Arismendi-Zambrano et al. 2021). In more recent times, papers have turned towards deep learning to address this problem, such as Gorodnichenko et al. (2021). Others have moved to measure the extent to which uncertainty is communicated in releases by both dictionary and non-dictionary approaches (Baker et al. 2016, Hassan et al. 2019, Caldara and Iacoviello 2022).

On the topic side, many studies have used approaches adapted from the Latent Dirichlet Allocation methods of Blei et al. (2003). Parle (2022) uses a dynamic adaptation of this algorithm to measure tone for each ECB president, while Cross and Greene (2020) use a non-negative matrix factorization approach. Other papers that examine topic include Hansen et al. (2019) for the Bank of England, Istrefi et al. (2021) for the Fed and Aguilar and Pérez-Cervantes (2022) for the Banco de México. Other studies using such methods for central banking analysis include Bybee et al. (2020), Hendry and Madeley (2010) and Boukus and Rosenberg (2006).

The time dimension is the least explored thus far in central bank literature. Thus far we are aware of a small number of papers that examine this dimension such as the work of Galardo and Guerrieri (2017), whereby the number of uses of "may" or "might" in a text are counted, as these are seen to have an unambiguous forward looking component. Coenen et al. (2017) look at the forward guidance component of communication to show that such communication reduces uncertainty, provided it is state contingent or providing guidance about a long period of time. Angelico et al. (2022) identify future-oriented tweets in Italian using future tense, categorical adjectives indicating the future, or both.

3.2 Our ECB and Fed Textual Data

In this study we examine textual data from two central banks, the US Federal Reserve (Fed) and the European Central Bank (ECB). For both cases, we use textual data from different forms of communication at different points in our study. Complete summary statistics for the textual data used are presented in Table 1.

For the Fed, we focus on three types of textual data: (1) announcements during policy event days; (2) the minutes of policy events, released after a policy event with a short lag; (3) textual data from the Greenbook accompanying files, released after a policy event with a significant time lag of 5 years.

The textual data relating to the policy event days can be sub-divided further. The Fed releases a short statement summarising the decision at the beginning of given policy release events. We have data relating to 207 statements in our dataset. Of these, 197 were scheduled meetings with 10 being unscheduled. Beginning in April 2011, the Fed introduced press conferences following policy meetings (at first for alternate meetings, and since 2019 for every meeting). We have 57 press conferences in our dataset. From these conferences, we have the transcripts of the President's introductory statement. We will always refer to this corpus as that of the "introductory statement", to distinguish it from the non-verbal "statement" that is released on the website. We also have transcripts of the questions and answers during the press conference.

We also have access to Greenbook textual data, which are derived from the supporting materials presented to the members of the FOMC prior to each meeting, and include not only charts and tables, but extensive verbal discussions.

For the ECB, we use the ECB press conferences (including the introductory statement, and the associated Q&A session with journalists) and the text of speeches delivered by members of the ECB executive board (discussed in section 5). We analyse the press conferences starting with the first conference of 9th June 1998 and ending with the conference of 10th September 2020. We have a corpus of 240 introductory statements by the ECB President, and we also have 237 Q&A sessions. [10] We also source data from speeches by ECB Executive Board members, which we extract from the ECB Speeches dataset, that can be found on the website of the ECB.

When conducting event studies of policy meeting days, we will consider all textual data released within the window of the event as one combined corpus. This means, for the case of the Fed, we amalgamate textual data from the statement, the introductory statement to the press conference, and the answers to questions during the press conference, as well as any textual data relating to unscheduled conference calls. For the case of the ECB, we combine textual data from the introductory statement to the press conference, as well as the answers to journalist's questions. We refer to the combined corpora as the policy event corpora. Note that we do not include the questions of the journalists themselves in our event corpora, though these questions do occur during the event windows.

In our investigations, we will map textual data from the policy event corpora to high-

[10] The data from the ECB press conferences were manually taken from the website of the ECB: `https://www.ecb.europa.eu/press/pressconf/html/index.en.html`. The number of Q&A sessions is lower than the number of statements because questions were not taken during the first three press conferences.

frequency asset price movements in response to the announcements. For the case of the ECB, we have textual data from the introductory statements and the press conference answers at a regular frequency throughout the sample, since the ECB conducted press conferences from its inception. For the case of the Fed, however, there is greater sparsity of textual data across the sample period. Firstly, for the period in the 1990s, there are decision days for which the FOMC made no statement, meaning there was no textual data. Secondly, the press conferences began in 2011, meaning that we have greater volumes of textual data for policy events after this date.

Though the sources of data present the text in a relatively regular manner, in order to use these data we apply some standard cleaning procedures. One important difference is that we need to preserve numerical information associated with dates, whereas numerical information is often jettisoned in other applied computational linguistics studies. A full discussion of cleaning procedures is detailed in Byrne et al. (2023).

Table 1: Corpora Characteristics

	Documents	Start	End	Median	Std	Min	Max
Fed Statements	207	1994/02/04	2021/12/15	15	6.09	4	33
Fed PC: Introductory Statement	57	2011/04/27	2021/12/15	61	14.00	23	113
Fed PC: Answers	57	2011/04/27	2021/12/15	226	86.33	75	539
Fed PC: Questions	57	2011/04/27	2021/12/15	110	34.26	48	266
Fed Minutes	413	1976/01/20	2021/12/15	196	150.37	3	834
Fed Greenbook: Part 1	145	1990/02/07	2010/04/28	997	138.24	551	1830
Fed Greenbook: Part 2	149	1990/02/07	2010/04/28	1777	387.95	34	2739
Fed Speeches	4568	1993/04/02	2022/05/11	118	70.51	2	1276
ECB Statements	240	1998/06/09	2020/09/10	60	15.09	16	133
ECB Answers	237	1998/10/13	2020/09/10	153	47.12	34	238
ECB Questions	237	1998/10/13	2020/09/10	63	16.08	23	116
ECB Speeches	2203	1997/02/07	2020/09/15	112	69.38	4	793

Notes: This table shows key summary statistics regarding the number of documents in each corpora and sub-corpora used in this study. The table also displays information on the number of sentences per document.

3.3 Algorithm: Temporal Tagging

The approach to temporal tagging used in this study categorizes textual data according to three dimensions:

1. *categorical* references to time: these refer to time only in a general sense, and include references such as "the future", "in the long-run", "currently";

2. *numerical* references to time: references that can be placed on a calendar such as "next year", "in the last few months", as well as more direct numerical references such as "1st January 2020" and "2020";

3. *grammatical* tense: our tagging algorithm also recognises whether sentences include the use of the present, past, and future tenses.

We shall now explain the tools we used to tag categorical, numerical, and grammatical references in sequence. There is an accompanying technical guide to the algorithms which will provide the interested reader with even more information (Byrne et al. 2023).

To extract numerical and categorical temporal references, we use the SUTime temporal tagger created by Chang and Manning (2012). SUTime is a rules-based temporal tagger built on regular expression patterns rather than on statistical relationships. SUTime is able to extract not only explicit references to dates ("January 2020"), but also relative date formats such as "two months from now", conditional on a reference date for a given document.

We make a number of adjustments to the SUTime algorithm, to take into account the particularities of central bank communication. For example, central bank policymakers will often use phrases such as "short-term", or "long-run" in their discussions, referring to periods in the future. We add such phrases to the SUTime library, since these were not previously included. We also created a list of commonly understood historical events, that often feature in our textual data. The speaker may refer to the "Great Depression", "Bretton Woods era" or "Global Financial Crisis" for example, and we tag such references by their underlying dates. These adjustments are described in in Byrne et al. (2023).

To tag usage of the past, present, and future tenses, we use the Tense-Mood-Voice (TMV) tool of Ramm et al. (2017). The TMV algorithm uses a list of rules to classify verbal complexes (sequences of verbal tokens within a given verbal phrase) into their tense. The system takes as its argument textual data to which POS tags have already been assigned.[11] The TMV algorithm is designed to isolate verbal complexes from the parsed sentences, before applying a rich sequence of rules to identify their tense[12]. For example, TMV will assign the verbal complex "I will go" to the simple future tense, since it recognises that the modal auxiliary "will" precedes the infinitive form of the verb ("go"). TMV recognises four forms of each of the past, present, future tense and two each of the conditional past and future tenses. We consolidate the TMV output into three general tenses: past, present and future, incorporating the conditional forms.

[11]Ramm et al. (2017) first parse textual data using the MATE parser (Björkelund et al. 2010), and we use the same algorithm.

[12]Note that although English does not have an inflectional future tense, as is found in some other languages, the "future tense" is an output of the TMV rules-based approach. We follow Ramm et al. (2017) in referring to the "future tense" throughout.

Table 2: Example of Our Temporal Parsing Approach

Past Tag	Textual Data	Future Tag
Tense	*"In the absence of improvement, such that the sustained return of inflation to our aim is threatened, additional stimulus will be required. In our recent deliberations, the members of the Governing Council expressed their conviction in pursuing our aim of inflation close to 2% in a symmetric fashion.* ECB President Draghi, Speech at Sintra, 18th June 2019.	Tense
Numerical, Numerical	*"Over the medium term underlying inflation is expected to increase, supported by our monetary policy measures, the ongoing economic expansion and robust wage growth. . . . This assessment is also broadly reflected in the September 2019 ECB staff macroeconomic projections for the euro area, which foresee annual HICP inflation at 1.2% in 2019, 1.0% in 2020 and 1.5% in 2021."* ECB President Draghi, Press Conference, 12th September 2019.	Categorical, Tense*, Tense*, Numerical, Numerical

Notes: Phrases marked `Numerical` are tagged as future/past using the SUTime tool. Phrases marked `Categorical` are tagged using the SUTime tool, with an additional bespoke dictionary of central-banking-specific future words (for example, "medium-run"). Phrases marked with `Tense` are tagged as future/past tense using the TMV tool. Phrases marked with `Tense*` are tagged as present tense using the TMV tool, but coded as future using a bespoke dictionary of present tense phrases that evoke future considerations, designed for use with central bank communications (for example, "expect", "forsee").

Having assigned verbal complexes to their tense, we applied an additional adjustment. In the language of central bank communication, we frequently observe statements such as "we expect", "we forecast", or "we predict". While these verbal complexes are in the present tense, and will be identified as such by TMV, they typically express views about the future. We therefore re-assigned such verbal complexes from the present to the future tense. The full list of verb forms that we re-assign are reported in Byrne et al. (2023), along with more detail on the TMV algorithm.

To recap, by applying the SUTime and TMV approaches, we identify numerical and categorical time-references (SUTime), as well as past, present, and future verbal complexes (TMV). Table 2 shows two example sentences from our corpora. The blue highlighted text captures future tagged content while the orange highlights capture the past references. Phrases marked `Numerical` or `Categorical` are tagged using the SUTime tool (or our additional central bank time references), while phrases marked with `Tense` are tagged using the TMV tool. Phrases marked with `Tense*` are tagged as present tense using the TMV tool, but coded as future using our bespoke dictionary of present tense phrases that evoke future considerations.

The first sentence from the speech at Sintra is clearly about the future, and the second is clearly referring to the past but providing context to their recent decision making. Our argument is that such context, especially the symmetric nature of their objective function, is useful and important information for markets trying to predict future interest rates.

The sentences from the Press Conference highlight another important aspect of the measures. The first of these sentences shows that we capture important forward-looking information with the adjusted tense reference as well as the categorical SUTime measure; standard tense analysis would either miss this reference, or even classify it as past tense. The second of these sentences highlights that the numerical references may sometimes blur the measure of temporal orientation coming from the other measures. Our baseline measures described below classify future and past according to sentences tagged by *any* of the measures. We also conduct our analysis using topics constructed according to each temporal tagger separately; this more disaggregated approach will generally generate even stronger results though the qualitative nature of the baseline results is the same.

3.4 Measures of Document Temporal Orientation

Once sentences are tagged appropriately, we are in a position to create measures of temporal orientation from our textual data. Our aim here is to develop document-level measures of past and future orientation. Here, future (past) orientation is understood as cases in which a given document (or group of documents) contains a large or small amount of future (past) information. While we focus on overall temporal measures initially, in the following section we will construct similar measures that are disaggregated by topic. To do this we first identify all sentences in our corpora that contain *at least one* reference to the future, be it a numerical future reference, a categorical future reference, or the use of the future tense. We then identify all sentences in our corpora that contain at least one reference to the past (according to any form of tag).

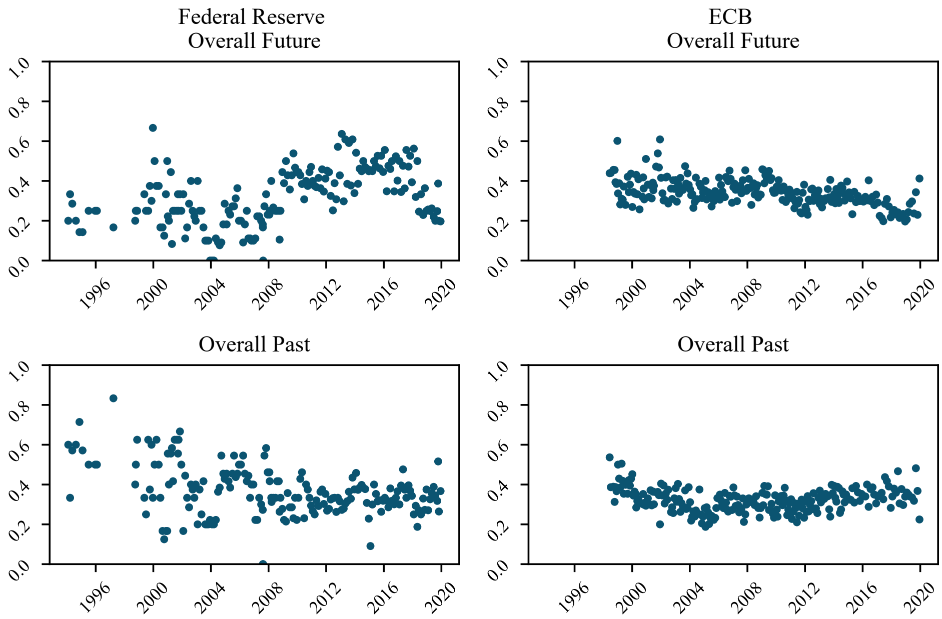
Formally, consider a sentence $s \in \{1, 2, \ldots, N^d\}$ found in document d. Let T_{ds}^{fut} denote a dummy variable taking value 1 if and only if sentence s contains at least one temporal expression relating to the future (be it numerical, categorical, hor the use of the future tense). Define T_{ds}^{pst} analogously as an indicator variable that equals 1, if and only if sentence s contains at least one reference to the past. Note that a sentence can be tagged as *both* about the future, and about the past, according to this scheme.

The document level overall temporal orientation measures, p_d^{fut} for future and p_d^{pst} for past, are defined as:

$$p_d^{fut} = \frac{1}{N^d} \sum_{s=1}^{s=N^d} T_{ds}^{fut}, \quad \text{and} \quad p_d^{pst} = \frac{1}{N^d} \sum_{s=1}^{s=N^d} T_{ds}^{pst}.$$

These measures can be interpreted simply as the fraction of sentences in a given document with at least one future or past tag, respectively.

Figure 2: The Time-Series of Overall Temporal Orientation of Policy Events



Notes: Figure shows the overall fraction of sentences marked with at least one temporal tag (be it future or past, respectively). The textual data are derived from the policy event corpora. The textual data from the Federal Reserve include statements (scheduled and unscheduled), introductory statements to press conferences, and answers. The ECB textual data contain information from introductory statements to press conferences, and answers. Textual data from event days are collated together when constructing these measures.

Figure 2 shows the time-series of the future and past overall orientation measures, for both the Fed and ECB. Note that the share of sentences marked as future and past fluctuates around 40% for each of the corpora. The future and past shares do appear to be inversely correlated for particular periods, which makes sense, given their construction. The negative correlation is not a strong one, however, taking values of -0.11 for the Fed, and -0.06 for the ECB. We have constructed our temporal taggers to allocate phrases or words to the mutually exclusive categories of past, present, and future. Ceteris paribus, an increased number of sentences tagged only as future will lead to a reduced past share. However, it is still possible that the future and past shares can co-move positively if the share of sentences marked only as present falls.

From Figure 2, several stylised facts emerge about temporal information in central bank communication. For the case of the Fed, we observe an increase in the future orientation of policy event communication from 2009 period to around 2014. Interestingly, this broadly overlaps with the forward guidance period. This evidence is also consistent with the evidence for the Fed statements presented in Coenen et al. (2017). We do observe, however, a fall in future orientation after around 2018. For the ECB, we observe a gradual secular decline in the overall future orientation, despite the fact our sample includes a period of forward guidance (from July 2013 onwards).[13] Our overall estimates therefore lend limited support for the finding of Coenen et al. (2017) that "central bank statements have on average become more forward looking".

3.5 A Validation Exercise for Document Temporal Orientation

At this stage we will establish a relationship between the temporal indicators ("futureness" and "pastness") and the distinction between an evaluation step and a projection step, as detailed in Section 2. We find it reasonable to believe that communications regarding evaluation will include more frequent references to past data, while communications regarding projection will include more future references. Of course, however, given the complexity of language, any given statement regarding evaluation or projection is likely to include both references to past and future. In this study we posit that evaluation and projection can be distinguished by their average levels of past and future references. It is on this basis that we interpret later results indicating the importance of future and past temporal references at explaining high-frequency news.

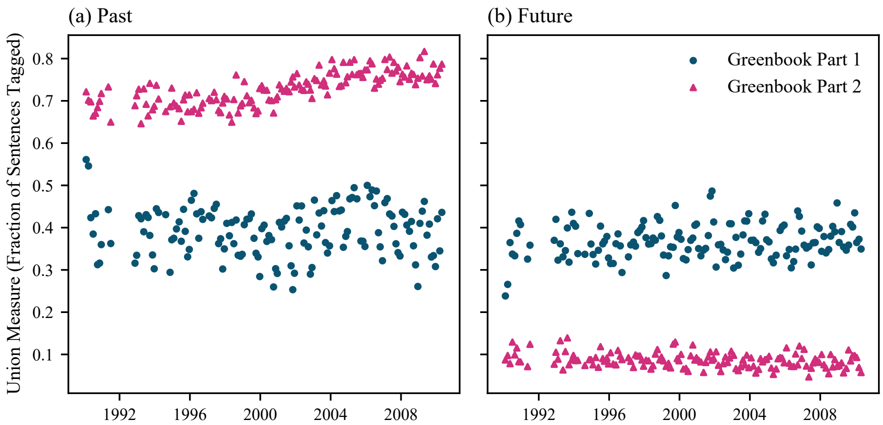
In this sub-section we do however propose an empirical validation exercise for our proposed relationship between the temporal tags generated by our algorithm, and the evaluation and projection steps. To do this we use data from the Fed. Specifically, we use the Greenbook data, which are a collection of discussions regarding the economic situation as well as forecasting exercises, and are produced to accompany each policy meeting of the FOMC. For a portion of our sample, the Greenbook textual data are actually divided into two distinct parts by Fed economists. The first part relates specifically to the forecasting exercises. The second part represents a series of evaluations regarding the present state of important economic variables, as well as recent trends. Importantly, the division of the corpora into two was performed by Fed staff, and is not a subjective division by

[13]Coenen et al. (2017) were not in a position to observe the falls in future orientation for the Fed post 2018, since their sample ended in March 2017. The behaviour we observe for the ECB is not evident in the estimates of Coenen et al. (2017) based on introductory statements. Note that these authors use the algorithm of Galardo and Guerrieri (2017), that does not identify numerical or categorical references, and uses a simplified approach to grammatical tense.

the authors of the present study. Moreover, the fact that the Fed divided their analyses into separate evaluation and projection portions lends weight to our assertion that the distinction between evaluation and projection is of important practical relevance for the conduct of central banking.

The two parts of the Greenbook data represent ideal corpora for evaluation, since we would expect our document level measures of futureness to be higher when applied to part one, relative to part two. Conversely, we would expect our document level measures of pastness to be higher when applied to part two, relative to part one. This is exactly the relationship that we find, as demonstrated in Figure 3. The measures generated by our algorithm are entirely successful at distinguishing between the evaluation portion of the Greenbook corpus, and the projection portion of the Greenbook corpus. For no meeting does part two of the corpus have a lower measure of pastness than part one, and the converse is true for the measure of futureness. This exercise therefore provides important direct evidence that our interpretation of futureness and pastness as measures of evaluation and projection is a reasonable one to make.

Figure 3: Validation of Our Temporal Indicators



Notes: Figure shows temporal measures constructed based on parts one and two of the Greenbook materials. Panel (a) displays the fraction of sentences tagged with at least one past tag. Panel (b) displays the fraction of sentences tagged with at least one future tag.

3.6 Construction of Temporal Topics

While these document-level future and past orientation measures may provide reasonable summary measures of the overall temporal orientation of a speech, or press conference, one concern is that such an approach would amalgamate information from a fairly diverse range of subjects. It is also of interest to study whether the temporal orientation of given topics within a speech changes the way market participants respond to given speeches. For example, a discussion of the future path of interest rates in a given speech may be more relevant for market participants than a discussion of past interest rate choices (even though the context of how the past is shaping current decisions may also be informative).

We shall assume that sentences are generated according to the following process. Each word is a random draw from the overall vocabulary, and the probability of drawing a given word from the vocabulary is determined by its topic. The topic of the word is itself randomly drawn from a sentence-level distribution over topics. A topic is thus synonymous with a probability distribution over words, and each sentence is associated with a probability distribution over topics. Let $\theta_{k,ds}$ denote the probability associated with topic k, $k \in (1, \ldots, K)$ in sentence s and document d.

Expressed in words, the future (past) temporal topic for a given document represents the average topic share for that topic, when we restrict the corpus to only those sentences that contain a future (past) reference. Formally, for a given document d with N_d sentences, and a given topic k, $k \in (1, \ldots, K)$ with $K = 15$, we denote the statement (or speech) future oriented topic measure ($\theta_d^{fut}(k)$) and the past orientation measure ($\theta_d^{pst}(k)$) respectively according to:

$$\theta_d^{fut}(k) = \frac{1}{N_d^{fut}} \sum_{s=1}^{s=N_d} T_{ds}^{fut}\, \theta_{k,ds}, \quad \text{where } N_d^{fut} \equiv \sum_{i=1}^{i=N_d} T_{di}^{fut},$$
$$\theta_d^{pst}(k) = \frac{1}{N_d^{pst}} \sum_{s=1}^{s=N_d} T_{ds}^{pst}\, \theta_{k,ds}, \quad \text{where } N_d^{pst} \equiv \sum_{i=1}^{i=N_d} T_{di}^{pst}.$$

Of course, the parameters $\theta_{k,ds}$ are not observable, and must be estimated. To measure topics, we use Latent Dirichlet Allocation (LDA) following Blei et al. (2003) and applied to central bank communication in Hansen et al. (2017). We apply LDA at the sentence level (rather than at the document level). We chose to estimate distinct topic models for the ECB based on its policy event corpus. For the Fed, we estimate the topic model based on the policy event corpus in combination with the minutes. The reason is that the sparsity of statement data limits our ability to estimate an interpretable topic model for the Fed, so we bolster the textual data from the policy events with related information from the minutes, when estimating the topic model. This is the only point at which the minutes data are analysed in our study, and they are not used in event studies, since they are not released during the events.[14] The topic model for the Fed is estimated on textual

[14] A number of pre-processing steps are taken prior to the fitting of the topic model. Note that these pre-processing steps are applied to the corpus prior to the application of the LDA model, but not prior to the application of the SUTime and TMV tools. These steps are largely standard, and include the removal of numbers and punctuation, the removal of standard English stopwords (e.g. "the"), the conversion of words to lower case and the use of a stemmer to reduce words to their root form (so "developing" and

data from the decision of 4 February 1994 to that of 11 December 2019. The topic model for the ECB is estimated on textual data from 9 June 1998 to 10 September 2020.

We set the number of topics to 15, which led to a distinct and interpretable series of topics. In order to interpret the topics, we examine those words that are associated with the topics with highest probability. We display the ten most probable words, conditional on topic, in Tables A.2 and A.1. The fifteen generated topics each have meanings that are broadly interpretable in line with certain aspects of central bank communication. A number of topics are directly related to inflation (Topics 2 and 3 for the Fed, and Topics 3, 6, and 7 for the ECB), while other groups are more directly related to monetary policy actions (Topics 4 and 11 for the Fed, and Topics 14 and 15 for the ECB). Certain topics are closely associated with discussion of underlying data (Topics 1, 10, and 5 for the Fed, and Topics 5 and 10 for the ECB).

Having successfully estimated topic models at the sentence level, we are able to construct the temporal topic measures. To demonstrate some of the properties of our new measures of temporal orientation, Panel (a) of Figure 4 displays the evolution of measures relating to topic 3 for the Fed and topic 7 for the ECB over time. Both of these topics relate to inflation, as can be seen from the word clouds presented in Panels (b) and (c). These measures are constructed for the policy event corpora, and therefore vary at the frequency of the policy meetings.

Part of the contribution of this study is to separately estimate temporal topic shares, and their evolution is also plotted in Figure 4. We can observe several features of the shares immediately. The first is that they are correlated with the overall topic share. This is unsurprising, given that they are calculated essentially as "sub-corpora" of the overall corpus.[15] The measures also appear more volatile than the overall measure. Again, this was perhaps to be anticipated, given that they are averages applied to a smaller number of sentences (i.e., those only those sentences that are tagged as future or past, respectively). This makes it likely that the new measures exhibit a somewhat greater degree of noise. However, what is also clear from Figure 4 is that the new temporal topics are not reducible to the overall topic measure, and that they can display interesting dynamics that are less pronounced in the overall measure. For example, for the case of Fed topic 3, we observe large measures of the past-topic in the post-2012 period. In the case of Topic 7 for the ECB, we note that there is a clear increase in the future topic share during the 2008

"developed" are both stemmed to the same root, "develop"). The model was estimated using a Gibbs sampling approach, with a burn in of 1000 iterations and a total of 2000 iterations for the fitting process.

[15]For Fed topic 3, the correlation of the future temporal topic with the overall topic is 0.56, and the correlation of the past temporal topic with the overall topic is 0.58. For ECB topic 7, the correlation between the future topic share and the overall share is 0.9, while the correlation between the past share and the overall share is 0.71. The correlation between the past and future of Fed topic 7 is positive, at 0.37. The correlation between the future and past share of ECB topic 8 is also correlated, at 0.69.

crisis period. There is also a larger increase during the period of 2015 to 2019, during which the ECB was conducting purchases as part of its QE programme. This was also the period during which the ECB President linked net asset purchases to a "sustained adjustment in the path of inflation", meaning the future evolution of inflation received a great deal of emphasis in discussion.

4 Analysis of the Yield Curve News

The framework presented in Section 2 suggests that central bank communication can cause market news when market participants, in aggregate, update their beliefs because the communication addresses an information deficit that they had. While some of the information deficit may relate to the central bank's outlook for the economy (the projection dimension), and how they might react to future states, it may also relate to assessments of the state of the economy and the need for contextualisation of current data.

Armed with our measures of the temporal dimension of communication, we can provide an empirical assessment of the temporal drivers of market surprises. If the information deficit only derives from the projection dimension of policy making, the inclusion or exclusion of past information should have no bearing on the extent to which we can systematically explain the asset price news. On the other hand, if conjunctural assessment matters, it should, at least on average, result in past temporal topics being able to explain some of the asset price news.

4.1 Empirical Framework

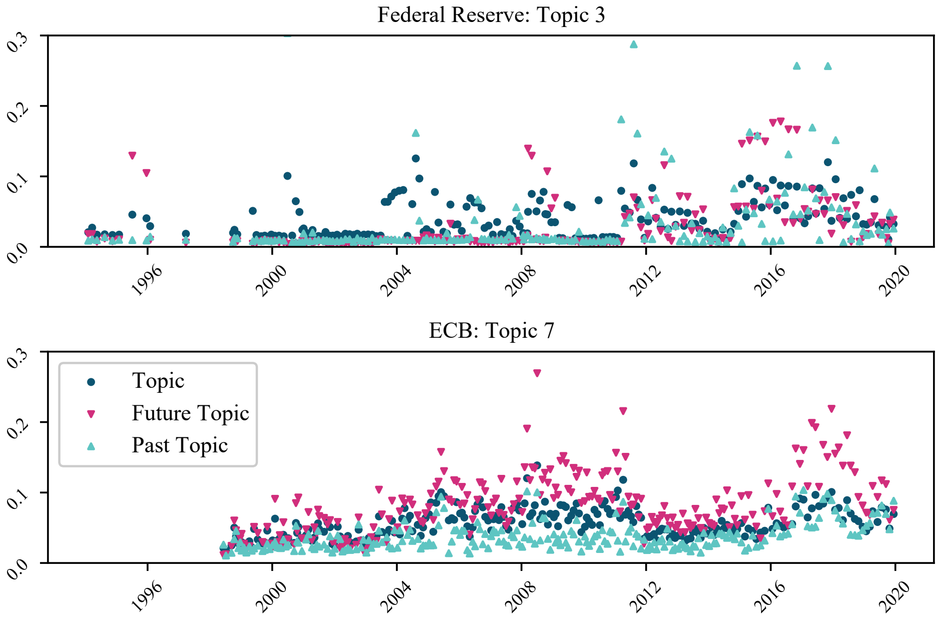
In theory, with a large enough data set, we could explore the empirical relationship between the market news and the temporal dimension of communication using a simple OLS regression of our market news variable on all the temporal dimensions of communication controlling for other released information such as numerical forecast information. In practice, we have a large number of independent variables of interest (our baseline specification has 14 topic main effects, 14 future topics, and 14 past topics, as well as other control variables) and relatively few observations.[16] Therefore, we follow Hansen et al. (2019) and adopt the "elastic net" LASSO specification of Zou and Hastie (2005).

Take a sample of N observations of a given response variable $\{y_i\}_{i=1}^{N}$ and a corresponding observations of a vector of p potential predictor variables $\{\mathbf{x}_i\}_{i=1}^{N}$, where $\mathbf{x_i}$ is

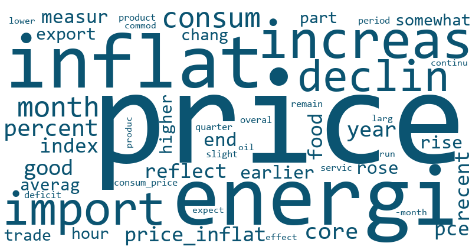
[16]The topic model contains 15 topics, but we drop one topic measure in regressions to avoid multicollinearity.

Figure 4: Topics and Temporal Topics, Two Examples from the Statements Corpus

(a) Evolution Over Time, Federal Reserve Topic 3 and ECB Topic 7



(b) Federal Reserve: Topic 3



(c) ECB: Topic 7

Notes: Panel (a) of the Figure displays the evolution over time of the document-level topic share, the future topic share, and the past topic share, for two example topics, topics 7 and 12. Note that these measures are derived from the statements corpus (including the introductory statement and the answers during the press conference). In the case that there is more than one statement per month, the values are averaged. Panels (b) and (c) of the Figure display, for reference, the estimated "word clouds" of topic 7 and 12 respectively, i.e. a representation of the highest probability words associated with the topic, where the size of the word indicates its weight.

of dimension $(p \times 1)$. The following minimisation problem is:

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{2} \sum_{i=1}^{N} (y_i - \mathbf{X}_i \beta)^2 + \lambda \left[\frac{1}{2}(1-\alpha)||\beta||_2^2 + \alpha||\beta||_1 \right] \right\},$$

for some $\lambda \geq 0$ and $\alpha \in [0,1]$, where $||u||_p \equiv \sum_{j=1}^{N} (|u_j|^p)^{1/p}$ is the l_1-norm. Here $y = (y_1, \ldots, y_N)$ denotes an N-vector of responses of interest, and $\mathbf{X}$ is a $N \times p$ matrix of independent variables. Note that when $\alpha = 1$ this specification is standard LASSO.

We define the parameter α to be equal to 0.99 and estimate the parameter λ by 10-fold cross-validation. We measure market news as the absolute value of the change in yields; this gives a measure of the update in beliefs without direction mattering.

One complicating factor when using LASSO estimation is that, in the case of highly correlated independent variables, parameters can be selected by the routine in a somewhat arbitrary manner.[17] To account for this, we summarise the results from our LASSO estimation routine using a non-parametric bootstrap. We draw with replacement from our dataset M observations, and do this 500 times, storing the distribution of estimated parameters. Note that for each bootstrapped dataset, we estimate a different value of λ via 10-fold CV. To compute the distribution of adjusted R^2 we adopt the following procedure. For a given bootstrapped dataset, having estimated a LASSO specification, we then re-estimate our prediction equation via OLS, conditional on the subset of variables that were assigned non-zero coefficients by the LASSO algorithm. Our measure of adjusted R^2 for this draw is taken from this OLS regression.

4.2 Asset Market Data and Control Variables

A growing literature in the empirical macro-economics literature focuses on the extraction of monetary policy surprises from policy announcements, by isolating changes in asset prices in narrow windows around the announcement. Early papers studied unidimensional monetary policy surprises (Kuttner 2001), and thus to study the impact of our textual measures of Fed and ECB communication on market yields, we use intra-daily data from Bauer and Swanson (2022) for the Fed, and the Euro Area Monetary Policy Event Database (EA-MPD) of Altavilla et al. (2019) for the ECB. These data are the change in asset prices in response to the Fed and ECB statements and press conferences, recorded as the difference in the price of assets before and after a narrow window (around 30-45 minutes) of each press conference or announcement. The assumption is that mone-

[17]This issue is widely known in the LASSO literature. Taddy (2017) cautions that (for the case of LASSO) cross-validation "can lead to over-fit for unstable algorithms whose results change dramatically in response to data jitter". See also Gentzkow et al. (2019) for discussion of this issue.

tary policy news should be the most important source of variation for these asset prices, given the tightness of the time-span of the window. In this case, our measures of the signal financial markets receive (in response to information from the press conference) should not be systematically related to other signals (for example signals about aggregate demand that do not come from discussions within the press conference). For the ECB, we exploit the fact that the press conferences follows the release with a lag, and use the press conference window only from the EA-MPD. All textual data for the ECB used in the event study comes from the press conference.

In the case of the Fed, the contracts of interest are the first and fourth Eurodollar contracts (ED1 and ED4), alongside Treasury yields with maturities of two, five and ten years (US 2Y, US 5Y and US 10Y). For the ECB, the contracts examined are OIS EONIA swaps with one month, one year and two year maturities (OIS 1M, OIS 1Y and OIS 2Y), while for the longer end of the yield curve German Bunds with yields of five and ten years are used (DE 5Y, DE 10Y).

Recent contributions in the monetary event study literature have followed Gürkaynak et al. (2005) and sought to decompose asset price movements into multiple forms of surprise, according to structural criteria. This study employs a number of monetary policy surprise series from leading recent papers in the literature, and this is discussed in more detail alongside results in Section 4.6.

Each LASSO specification contains a baseline set of controls. For the US, we use the 6 control variables used in the study of Bauer and Swanson (2022). These authors argue that the fact that market participants do not know the reaction function of the policymaker can lead to ex post predictability of surprises, given past financial data. For this reason these authors argue one should control for these financial data. For the ECB cases, we merely control for releases relating to the macroeconomic projections, which is numerical information released during the event window and could thus be correlated with both the dependent variable and our textual measures. For each of these sets of control variables, we force LASSO to include all such variables across all bootstrap draws via penalisation.

Finally, we should be clear on our sample selection. Since the Fed only began releasing statements in 1994, our event study for the US begins with the decision of 4 February 1994. Since Altavilla et al. (2019) state that the OIS data are too illiquid to analyse prior to 2002, we commence our event study for the ECB from the conference of 3 January 2002. We also exclude the pandemic period from our sample period, ending our event study with the Fed and ECB meetings of 11 and 12 December 2019 respectively. Analysis of the pandemic period is left for future work, in particular since the manner of communication shifted from in-person press conferences, to online video conferences. This could have introduced an important structural break in our textual data.

4.3 Effects of Temporal Communication on Market Yield News

Table 3 presents the bootstrapped mean Adjusted R^2 values from post-LASSO OLS regressions of the absolute value of changes in yields on sets of variables selected by the LASSO estimation. Results are shown for both the Federal Reserve (panel a) and the ECB (panel b). The main message is that the temporal dimension is highly relevant for explaining the news and both future and past temporal topics matter. This suggests that communication of both evaluation and projection are important for addressing the market's information deficit.

The most basic measures of communication content, the topics θ_k, capture some of the news systematically. Including the topics improves the Adjusted R^2 for the Fed sample by between 0.14 and 0.22 and for the ECB sample by between 0.15 and 0.25 compared with just the baseline controls for each case. This suggests that simply knowing what the press conference discussion is about is useful for picking up systematic variation in the market reaction.

Separately adding past and future topics to the independent variable set also increases explanatory power (rows 3 and 4). For the 2 year Treasury bond yields, the explanatory power increases from 0.21 on Adjusted R^2 to 0.3 (past added) and 0.29 (future added). For the ECB sample explanatory power goes from 0.28 on Adjusted R^2 to 0.36 (past added) or 0.37 (future added) for the 2 year OIS. In both cases, nearer the short end, the past topics actually capture more of the variation whereas further out the yield curve, the future topics typically explain more.

For both samples and across all asset classes, the strongest explanatory power comes from specifications that allow the LASSO to select over topics from both past and future dimensions. These additional increases in Adjusted R^2 suggest that the past and future topics are not measuring the same source of additional information, since using the measures in combination explains more news than specifications using only one set of measures.

Table 4 repeats the analysis but instead reports the Pseudo Out-of-Sample (OOS) Mean Squared Error (MSE). In this case, results are always compared to the MSE from a specification only with topics data. This MSE is then indexed to 100. A lower MSE arising from adding temporal indicators indicates that the addition of these variables improves the fit. The conclusions reached with the analysis of the Adjusted R^2 are endorsed with the MSE analysis in that our temporal measures explain news beyond what has been achieved in the previous literature.

4.4 Robustness of the Main Results

We perform two forms of robustness check on our baseline results for the yield curve. Firstly, we implement a placebo test to assess the reliability of our approach to mapping news signals to asset price responses. Secondly, we incorporate additional plausible control variables, based on the existing literature.

4.4.1 Placebo Tests

One immediate concern may be that, despite using the elastic net procedure to select only covariates that systematically help to explain the market news, we may nonetheless be achieving greater explanatory power simply by adding more disaggregate variables that the elastic net gets to choose from. To allay these concerns, we conduct a placebo test by randomly shuffling (with replacement) the generated topic and temporal measures across meetings, without permuting the controls and dependent variables. We then repeat the exercise of the previous sub-section on the newly created placebo dataset.
Table 5 is the equivalent of Table 3 but reports the results of this exercise. What is clear from these results, and was found in Hansen et al. (2019), is that the elastic net procedure does not naturally generate the explanatory power. The bootstrapped mean Adjusted R^2 values from the placebo test indicate essentially no improvement of fit from adding the placebo data. This is what one should expect from randomised data. We find similar effects looking at MSE but do not report the results here in the interests of space.

4.4.2 Additional Controls

In our study we began from the observation that asset price reactions to central bank statements imply the existence of an information deficit between market participants and monetary policymakers. We have found that temporal information accounts for an important portion of this deficit. One general concern is that our measures of temporality may correlate with other features of the textual data that could drive the results. In this subsection we investigate the roles of measures of uncertainty and of document tone, the second "T" of text. We first describe the additional controls we use, before discussing the results for each investigation.

Many of the early applications of NLP methodologies to finance found that the tone of information releases could influence asset prices (e.g., Das and Chen 2007, Tetlock 2007). To assess whether the tone of the documents within the event corpora could account for our findings, we construct a measure of the tone of given communication events (at the policy day frequency). We measure the tone of each document using the expansionary (i.e., positive terms) and contractionary economic terms (i.e., negative

Table 3: Adjusted R^2, Yield Curve News, Intraday Dependent Variables

(a) Federal Reserve

Specification	ED1	ED4	US 2Y	US 5Y	US 10Y
Controls Only	0.182	0.102	0.067	0.026	0.039
Topics	0.344	0.324	0.213	0.166	0.191
Topics, Past Topics	0.392	0.394	0.298	0.232	0.254
Topics, Future Topics	0.386	0.391	0.288	0.259	0.272
Topics, Future and Past Topics	0.432	0.462	0.379	0.318	0.327

(b) ECB

Specification	OIS 1M	OIS 1Y	OIS 2Y	DE 5Y	DE 10Y
Controls Only	0.030	0.072	0.072	0.008	0.014
Topics	0.256	0.288	0.282	0.187	0.150
Topics, Past Topics	0.356	0.368	0.357	0.220	0.194
Topics, Future Topics	0.322	0.362	0.368	0.251	0.195
Topics, Future and Past Topics	0.409	0.430	0.427	0.274	0.236

Notes: This table shows the adjusted R^2 for specifications predicting the responses of the absolute value of intraday changes in yields at different maturities for the Federal Reserve (top panel, a) and the ECB (lower panel, b). The "Controls Only" specification contains only the controls of Bauer and Swanson (2022) as independent variables, and is estimated by OLS. The remaining four specifications are estimated using an Elastic Net specification, which selects over different bundles of potential predictors. The "Topics Only" bundle contains controls, plus 14 topic main effects. The "Topics and Past" and "Topics and Future" specifications contain the same bundle as that for "Topics Only", and additionally allow for the inclusion of 14 temporal topics, respectively relating to past and future. The "Topics, Future and Past" uses the super-set of the previous bundles as the set of potential predictors. For the Elastic Net specifications we display the median estimated adjusted R^2 across 500 draws from a non-parametric bootstrap algorithm. For each bootstrap draw, the Elastic Net is estimated by 20-fold cross-validation. The adjusted R^2 is computed on the basis of OLS estimates of the prediction equation, using only those explanatory variables selected by the Elastic Net. For each Elastic Net specification, we force estimates to include the baseline control variables via penalisation. Intraday asset price movements for the Federal Reserve are taken from the dataset of Bauer and Swanson (2022). Intraday asset price movements for the ECB are taken from the dataset of Altavilla et al. (2019).

Table 4: Pseudo-OOS MSE, Yield Curve News, Intraday Dependent Variables

(a) Federal Reserve

Specification	ED1	ED4	US 2Y	US 5Y	US 10Y
Topics	100.00	100.00	100.00	100.00	100.00
Topics, Past Topics	96.78	97.75	96.58	98.40	98.70
Topics, Future Topics	100.44	98.07	98.17	97.02	97.17
Topics, Future and Past Topics	96.01	95.54	94.02	95.48	96.32

(b) ECB

Specification	OIS 1M	OIS 1Y	OIS 2Y	DE 5Y	DE 10Y
Topics	100.00	100.00	100.00	100.00	100.00
Topics, Past Topics	93.01	94.95	95.88	99.76	99.19
Topics, Future Topics	97.16	95.53	93.99	97.03	98.39
Topics, Future and Past Topics	90.68	90.74	91.06	97.50	97.55

Notes: This table shows the pseudo out-of-sample (OOS) mean squared error (MSE) for specifications predicting the responses of the absolute value of intraday changes in yields at different maturities, respectively for the Federal Reserve (top panel) and the ECB (lower panel). We express the OOS MSE relative to that of the "Topics Only" specification, which is normalised to equal 100. For each bootstrap draw, the pseudo OOS MSE is computed as the average MSE associated with the left-out fold of the 20-fold cross-validation procedure, under the optimal estimate of λ, across each of the 20 folds. We take the median OOS MSE across the 500 non-parametric bootstrap draws. Intraday asset price movements for the Federal Reserve are taken from the dataset of Bauer and Swanson (2022). Intraday asset price movements for the ECB are taken from the dataset of Altavilla et al. (2019). See Notes to Table 3 for further details regarding the Elastic Net specifications.

Table 5: Placebo Test: Adjusted R^2, Yield Curve News, Intraday Dependent Variables

(a) Federal Reserve

Specification	ED1	ED4	US 2Y	US 5Y	US 10Y
Controls Only	0.182	0.102	0.067	0.026	0.039
Topics	0.192	0.115	0.078	0.040	0.053
Topics, Past Topics	0.197	0.120	0.082	0.045	0.059
Topics, Future Topics	0.196	0.120	0.081	0.043	0.057
Topics, Future and Past Topics	0.199	0.124	0.084	0.047	0.061

(b) ECB

Specification	OIS 1M	OIS 1Y	OIS 2Y	DE 5Y	DE 10Y
Controls Only	0.030	0.072	0.072	0.008	0.014
Topics	0.046	0.086	0.087	0.024	0.030
Topics, Past Topics	0.058	0.092	0.093	0.029	0.038
Topics, Future Topics	0.058	0.092	0.092	0.028	0.037
Topics, Future and Past Topics	0.068	0.095	0.096	0.032	0.042

Notes: This table shows the adjusted R^2 for specifications predicting the responses of the absolute value of intraday changes in yields at different maturities, respectively for the Federal Reserve (top panel) and the ECB (lower panel). Independent variables (excluding controls) were reshuffled randomly across meetings, in order to construct a placebo test, as described in the text. Intraday asset price movements for the Federal Reserve are taken from the dataset of Bauer and Swanson (2022). Intraday asset price movements for the ECB are taken from the dataset of Altavilla et al. (2019). See Notes to Table 3 for further details regarding the Elastic Net specifications.

terms) from Hansen and McMahon (2016). The tone of each document is defined by:

$$tone_d = \frac{N_{pos,d} - N_{neg,d}}{N_{pos,} + N_{neg,}}$$

where $N_{pos,d}$ is the count of the number of positive terms and $N_{neg,d}$ is the number of negative terms within each document d We enter this textual measure of tone as an additional control.

Another robustness check is to assess whether our measures are somehow proxying for uncertainty. For example, if policymakers were to release more temporal information at times of high uncertainty, this would bias our estimates. We have good reason to be interested in the relation between our results and uncertainty, since recent papers have shown non-linearities in the effects of surprises and the uncertainty prior to monetary policy meetings (Bundick et al. 2021, Bauer et al. 2021, De Pooter et al. 2021). In addition our dependent variables are defined in terms of absolute values which means that shocks in times of greater uncertainty will naturally lead to greater movements.

To control for this kind of uncertainty, we add a control for the lagged value of market-based uncertainty prior to the monetary policy announcement, using the VIX derived from US and German equity markets. We can think of such market uncertainty regarding the state of the economy as being a mixture of aleatoric, ontological and epistemic uncertainty; there is an element of randomness of outcomes to the current economic process (aleatoric), it is difficult to know the precise levels of concepts such as potential GDP due to the requirement of model assumptions (ontological) and data limitations mean we cannot fully assess how new information will play out (epistemic).

In addition, we note that information from policy statements can in itself signal the central banks' beliefs about broader policy uncertainty, as is found in Hansen et al. (2019). As such, independent of the level of market uncertainty prior to the meeting, communication of uncertainty can have an impact and textual controls can have an additional contribution. We control for this using the count of words from uncertainty subset of the dictionary of Loughran and McDonald (2011) within each document normalised by the length of the text.

Finally, in this sub-section we will examine whether the structural features of the Fed policy releases affect results. As previously mentioned, prior to the introduction of press conferences in 2011, the Fed released relatively short statements. For this reason, our measures of temporality could be relatively noisy for those events without press conferences. Therefore we perform an investigation with press conferences removed. We also perform an investigation where we remove the unscheduled meetings from our sample for the Fed.

When repeating the exercises with additional controls, the additional variable, like all control variables throughout, is constrained to be included in all specifications (it may not be excluded by the LASSO algorithm).

We display a summary of our estimates for two contracts, for each of the two central banks, in Figure 5. We choose to display results for ED4 and the Treasury 10-year rate for the Fed, and the one-year OIS and 10-year German yield for the ECB.[18] We have generally found that adding additional controls for the influence of the potential omitted variables we have discussed makes little difference to our findings. Adding a measure of tone appears to raise the explanatory power of all specifications by a small amount, without affecting the "ordering" over specifications. Similar findings are in evidence for specifications with the two uncertainty controls. We conclude that our measures of temporality are not reducible to tone, or to uncertainty measures.

We do see some evidence that the unscheduled meetings weaken the relation between our temporal topics and the measures of market news. This can be seen for both ED4 and the 10-year yield in Figure 5a, and is particularly true for the future temporal topics. One explanation is that important information regarding future outcomes is more typically conveyed in the scheduled meeting statements, while the unscheduled meetings focus on extraneous contemporary shocks. Further, removing the press conferences has little effect on the qualitative findings.

4.5 The Relation Between Explanatory Power, Temporality, and Topic

While our previous specifications shed light on the information gained by adding temporal and topical information, it is difficult to assess which kinds of information are driving these gains. This is partly due to a dimensionality issue - with a large number of topics, identification of precise marginal effects is difficult. In order to try to overcome this challenge, we construct meta topics, combining fourteen of the topics[19] based on their similarity according to the distance between the topic-word distribution of each topic, as measured using the Hellinger distance.[20] These are clustered using a dendogram as shown in Figures 6.

We repeat the Adjusted R^2 analysis for the Fed and ECB using the fourth eurodollar futures contract (ED4) and one year OIS rates (OIS 1Y) respectively. In each case, we

[18]Results for the other contracts are available on request, though conclusions are unchanged.

[19]One of the fifteen generated topics is omitted in each case, in line with the main specification.

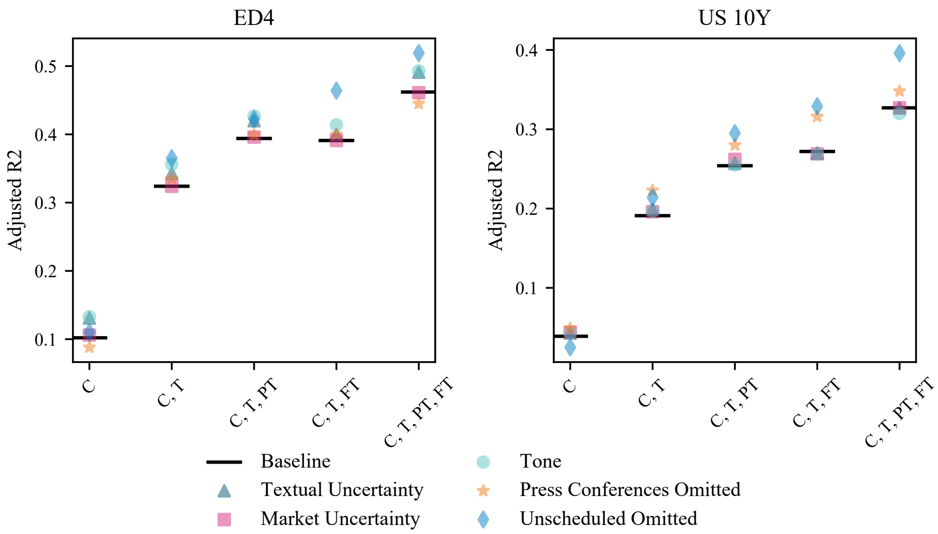
[20]Hellinger distance between topics A and B, where a_i and b_i is the weight on token i in each topic, is:

$$dist_{AB} = \frac{1}{\sqrt{2}}(\sum_{i=1}^{k}(\sqrt{a_i} - \sqrt{b_i})^2)$$

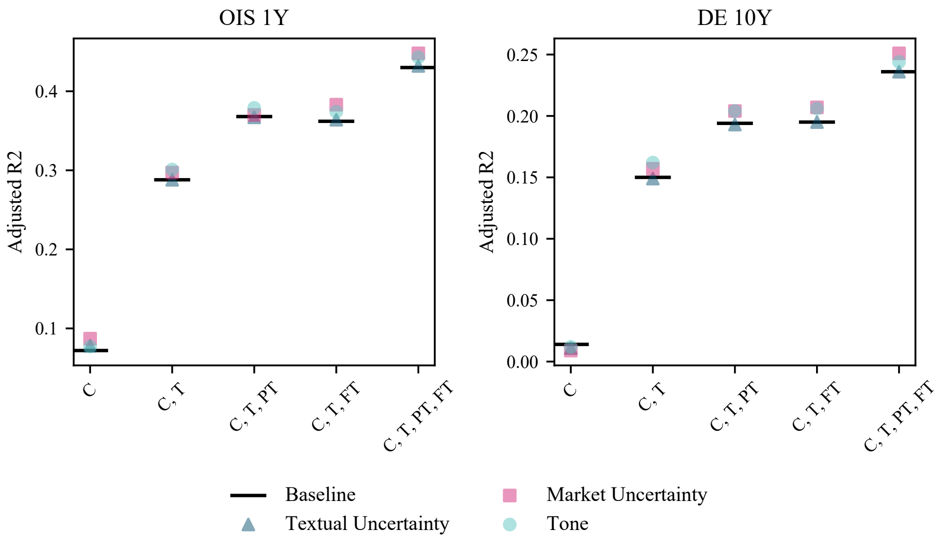
.

Figure 5: Summary of Robustness Exercises

(a) Federal Reserve

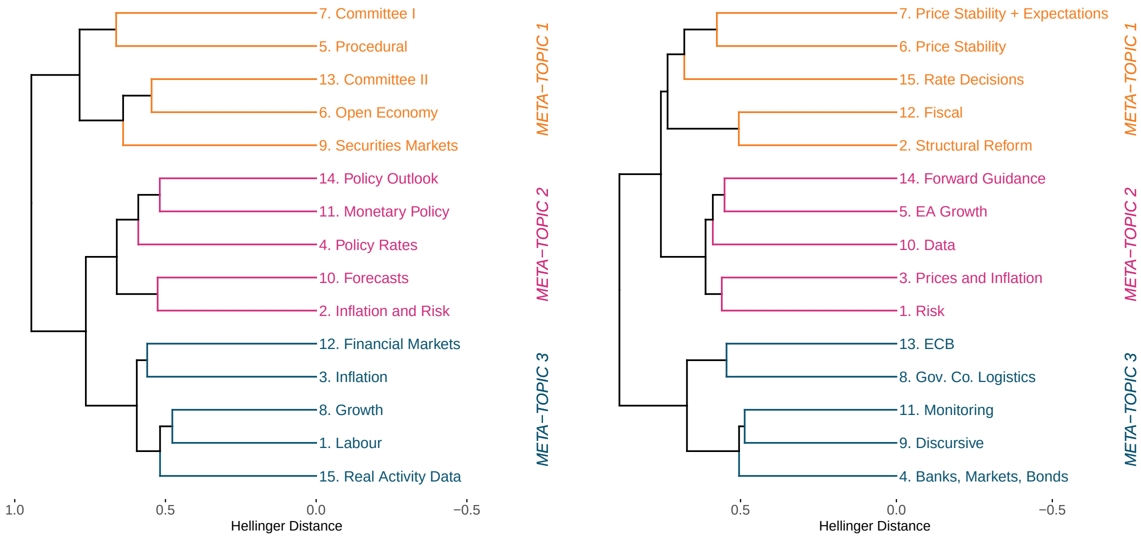


(b) ECB



<u>Notes:</u> This figure shows the adjusted R^2 for specifications predicting the responses of the absolute value of intraday changes in yields at different maturities, respectively for the Federal Reserve (top panel) and the ECB (lower panel). C: controls only; C: controls only; C, T: controls and topics; C, T, PT: controls, topics, past topics; C, T, FT: controls, topics, future topics; C, T, FT: controls, topics, future topics, past topics. In addition to the baseline controls, each specification contained additional different control variables: (1) a measure of textual uncertainty; (2) a measure of market uncertainty (lagged VIX); (3) a measure of the tone of the textual data. For the case of the Fed, we additionally estimate a specification with the press conferences omitted from the sample, and another with the unscheduled meetings removed from the sample. Intraday asset price movements for the Fed and ECB taken from the datasets of Bauer and Swanson (2022) and Altavilla et al. (2019) respectively. See Notes to Table 3 for further details regarding the Elastic Net specifications.

Figure 6: Dendogram clustering



Notes: This figure shows the dendogram constructed using Hellinger distance for the Fed topic model (LHS) and the ECB topic model (RHS). Each node shows the level of dissimilarity between each branch.

omit one meta topic each time, inferring that any reduction in the explanatory power can be attributed to removing this meta topic. The results are shown in Tables 6a and 6b.

For the Fed, the largest drop in Adjusted R^2 is seen when dropping Meta Topic 2, which is the topic most broadly related to monetary policy. This provides some evidence that topics more related to the monetary policy mandate more strongly drive the increase in explanatory power, both when included alone but also when interacted with the temporal aspects. In the case of the third meta topic, which relates to data analysis, there is not a large reduction in explanatory power in a topic specification. However the reduction is larger in the interaction with future temporal information, suggesting this meta topic may capture forecasting/projection information.

In the case of the ECB, a broadly similar narrative emerges. There is limited evidence that Meta Topic 1, relating to monetary policy, expectations and fiscal policy, has more explanatory power than the other meta topics, particularly when interacted with the temporal information. It should be noted however, that the explanatory power relating to broader economic discussions (Meta Topic 2) is also reasonably large.

On the whole, this evidence suggests that it is primarily monetary policy discussion, relating both to mandate and policy decisions, that may be driving effects, although there is also a role for broader economic analysis.

Table 6: Hellinger metatopics analysis

(a) Federal Reserve

Specification	Full	Cut MT 1	Cut MT 2	Cut MT 3
Controls Only	0.102			
Topics	0.324	0.278	0.261	0.305
Topics, Past Topics	0.394	0.366	0.304	0.370
Topics, Future Topics	0.391	0.341	0.290	0.337
Topics, Future and Past Topics	0.462	0.420	0.332	0.395

(b) ECB

Specification	Full	Cut MT 1	Cut MT 2	Cut MT 3
Controls Only	0.072			
Topics	0.288	0.232	0.234	0.256
Topics, Past Topics	0.368	0.278	0.300	0.332
Topics, Future Topics	0.362	0.279	0.285	0.335
Topics, Future and Past Topics	0.430	0.310	0.353	0.408

Notes: This table shows the adjusted R^2 for specifications predicting the responses of the absolute value of intraday changes in yields at different maturities, respectively for the Federal Reserve (top panel) and the ECB (lower panel). Specifications are identical to Table 3, with the exception that in each case one Hellinger distance generated meta topic is dropped (excluding the first baseline specification). Intraday asset price movements for the Federal Reserve are taken from the dataset of Bauer and Swanson (2022). Intraday asset price movements for the ECB are taken from the dataset of Altavilla et al. (2019). See Notes to Table 3 for further details regarding the Elastic Net specifications.

4.6 Monetary Policy Surprise Analysis

In our previous investigations, we examined only the response of asset prices across the yield curve to textual measures. In an influential contribution, Gürkaynak et al. (2005) argued that such asset price movements can be interpreted as reduced form responses to more than one monetary policy shock with a structural interpretation. These authors found that two such shocks were sufficient to explain observed variation. Therefore, an interesting question is whether our textual signals have differing capacity to explain the structural components of asset price responses.

Therefore we repeat our previous analysis, with structurally decomposed shocks as dependent variables. Specifically, we use the Swanson (2021) structural surprises for the Federal Reserve, which are measures of target, forward guidance, and large-scale asset purchase (LSAP) surprises. The study of Altavilla et al. (2019) applied a very similar decomposition to euro area data, generating target, forward guidance and quantitative easing (QE) surprises. These authors also extract an additional form of forward guidance surprise, which they term the timing surprise, since it relates to the timing of decisions, that are in some sense already "priced in" by market participants.

Neither of these two surprise series separates monetary policy surprises from so-called "information surprises". It is of particular interest to our study as to whether our textual data signals are related systematically to monetary policy or information effects. Many of the signals we have generated are direct measures of topics that are focussed on discussions of data, or forecasts. We have also motivated why discussions by policymakers of recent data may themselves be a source of news to market participants, to the extent they inform markets of central bank evaluation functions. For this reason we examine also the responses of information and monetary policy surprises based on the decomposition of Jarocinski and Karadi, in addition to the two decompositions already discussed.

We display our results in Table 7. The basic result of the earlier analysis remains - both forward and backward looking temporal topics play and important role in explaining news variation. For both the Fed and for the ECB, future temporal topics explain more variation in target surprise news relative to past temporal topics, though the difference in Adjusted R^2 is greater for the ECB. For the two forward guidance surprises, we also find that future topics have larger explanatory power than past topics, and again the difference is greater for the ECB. One explanation for the greater difference for the ECB would be that the decomposition of Altavilla et al. (2019) separately extracts timing surprises, for which past information has greater explanatory power than future. To the extent that this result is general, and the Swanson (2021) subsumes timing and forward guidance together, this may explain the stronger relation between future topics and forward guidance in the ECB cases. For the case of the Fed LSAP and ECB QE surprises, we only observe a difference between past and future topics for the case of the Fed, where past topics have greater explanatory power. This suggests that guidance regarding LSAPs was quite closely related to the evaluation step of the Fed.

Table 7: Adjusted R^2, Decomposed Surprises

(a) Federal Reserve

Specification	Target	FG	LSAP	INFO	MPOL
Controls Only	0.175	0.020	0.054	0.115	0.259
Topics	0.317	0.233	0.166	0.234	0.390
Topics, Past Topics	0.373	0.308	0.208	0.275	0.441
Topics, Future Topics	0.380	0.313	0.187	0.273	0.493
Topics, Future and Past Topics	0.431	0.373	0.228	0.309	0.526

(b) ECB

Specification	Target	Timing	FG	QE	INFO	MPOL
Controls Only	0.020	0.020	0.033	0.096	0.007	-0.023
Topics	0.217	0.280	0.289	0.208	0.165	0.167
Topics, Past Topics	0.275	0.360	0.329	0.262	0.262	0.238
Topics, Future Topics	0.330	0.352	0.342	0.263	0.271	0.251
Topics, Future and Past Topics	0.373	0.415	0.372	0.312	0.339	0.325

Notes: The top panel of this table shows the adjusted R^2 for specifications predicting the responses of the absolute value of intraday changes in (i) Swanson (2021) surprises at different maturities, (ii) Bauer and Swanson (2022) surprise (the first principal component, not orthogonalised), and (iii) Bu, Rogers, Wu (2020) surprise. The bottom panel of this table shows the adjusted R^2 for specifications predicting the responses of the absolute value of intraday changes in the Altavilla et al. (2019) surprises. See Notes to Table 3 for further details regarding the Elastic Net specifications.

Table 7 also shows the responses of the Jarociński and Karadi (2020) surprises by specification. As before, both temporal topics systematically contain information to explain each shock, and the joint inclusion captures the most variation. Here we observe some interesting differences between the Fed and ECB results. For the case of the Fed, our measures have consistently greater ability to explain the monetary policy surprise, and less so the information surprise. For the ECB, we observe a broadly comparable ability to explain the two surprises when we use the specification with both future and past topics. One explanation for this finding could be that the structure of Fed and ECB communication favours the mapping between textual data and information surprises for the ECB. The reason is that the volume of textual data revealed to market participants during the event window was far smaller, for the Fed, prior to the press conferences, as previously mentioned. Statements that were released before the period of press conferences were typically terse, and related only to monetary policy. On the contrary, the ECB conducted press conferences, during which macroeconomic data was discussed, throughout our sample period.

Another feature of our results is that we do not find strong evidence that the information effect is more related to the evaluation step (proxied by backward-looking information), relative to the projection step (proxied by forward-looking information), since both future and past topics relate to information shock news in a similar way across cases. We conclude that information effects are likely to reflect both projection and evaluation to a comparable degree.

5 Press questions as an information deficit measure

The analysis so far provides evidence that the central bank's conjunctural assessment is an important source of information for markets. However, it is difficult to specifically identify what the market does not know in order to assess the information deficit. To proceed, we develop an alternative, novel measure of the information deficit. This approach provides direct evidence of the time-varying nature of the deficit.

We exploit the Q&A part of the ECB press conferences and we focus on the questions asked by financial journalists as a measure of the information deficit. Between October 1998 and September 2022, there have been 6,175 questions asked by journalists when called upon by the ECB's Director General of Communications. Journalists typically ask two questions when they are called upon in the press conference. Our measure is based on the assumption that the questions highlight the issues that journalists, after hearing the opening statement, require clarification on, and that these reflect the information gaps that their financial market readers would also have. We also assume that clarification is sought on the most pressing information gaps first.

5.1 Validating the use of journalists' questions

While we argue that journalists' questions should be a signal of the information deficit at the time of the press conference, we wish to formally show this. For this, we extend our analysis to ECB Executive Board member speeches but combine the analysis with the questions from the preceding Q&A.

Our idea is as follows. Given that the impact of any communication is related to its newsworthiness, a speech which addresses market participants' information deficit is more likely to generate news and lead beliefs to update. The speech may clarify areas about which the market is unclear, or it may inform on the latest thinking of the central bank on the state of the economy or the reaction of monetary policy. In order to show that the questions reasonably reflect the information deficit, we shall show that speeches delivered in the inter-meeting period that address issues raised in the questions generate more market news than speeches that do not.

The speeches we use come from the ECB Speeches Dataset. This is an archive of all speeches by ECB Executive Board members, dating back to February 1997.[21] These data are continually updated by the ECB, we use the version of the dataset that ends with a speech on the 15th September 2020. Our sample includes 2,203 English speeches.[22] As more than one speech is often given on a particular day, the total number of days on which one or more speech occurs is 1,713. We treat speeches given on the same day essentially as one document because we will use a one-day asset price window. We otherwise process the speeches in the same way as the statements, though some more cleaning was required in the case of the speeches as discussed in Byrne et al. (2023).[23]

The temporal dimension of speeches generates similar, albeit less systematic, market news on the yield curve as the main announcements; Table A.5 in the Appendix presents the baseline yield curve analysis but applied to the speeches for both the ECB and Fed. This likely reflects that speeches are, on average, less central to market monetary policy belief formation. This may be because the speeches are not always about monetary policy. Or, even when they are about monetary policy, they are less likely to introduce new information and may often just restate the opening statement that has already been communicated (stale information may not get the same amount of attention Ehrmann and Sondermann (2012)). Nonetheless, we know that some speeches, such as Mario Draghi's "whatever it takes" speech have had large scale market impacts. Further to this, speeches can play a role in clarifying comments made in the previous press conference, or indeed can be used to steer the conversation ahead of upcoming monetary policy decisions.

In order to show that journalists' questions are an indicator of the information deficit,

[21]The dataset includes speeches delivered by senior officials prior to the formal creation of the ECB in June 1998. It is available here: `https://www.ecb.europa.eu/press/key/html/downloads.en.html`

[22]We exclude 159 non-English speeches. There are 16 speeches for which no text is available, which are discarded. There are 34 speeches that merely summarise the title and topic of the speech, and provide a hyperlink to lecture slides – these are also discarded.

[23]Note that our topic models were estimated on the policy event corpora, and therefore did not include data from the speeches. Given the event corpora topic models, we extrapolate our topic model to the textual data from the speeches (*querying* in the language of information retrieval). We prefer to use topic models that were estimated on textual data purely focused on monetary policy and its interpretation, rather than models that would incorporate information from the speeches that is frequently quite varied (including discussions of politics or law, for instance). This decision ensures consistency of topics across our investigations. The baseline fitted 15-topic LDA model was extrapolated out of sample to the relevant speeches dataset. A potential concern is the existence of words that may be in the speeches data but not the press conferences. Reassuringly, in total these account for only 4.3% of the total count of words in the ECB speeches that occur in the speeches dataset, and as such their exclusion from the training data is unlikely to have a fundamental impact on our goodness of fit.

we now identify the similarity of speech content and the questions (representing the information deficit). We compute our similarity measure using the document term matrix in which terms are weighted by the term frequency-inverse document frequency (TF-IDF) score.[24] Using this weighted document-term matrix, and letting dtm_i be a vector corresponding to the row of the weighted document term matrix associated with document i, we calculate the similarity between two documents i and j as:

$$Sim_{i,j} \equiv \frac{\sum_{k=1}^{n}(dtm_{ik} \times dtm_{jk})}{\sqrt{\sum_{k=1}^{n} dtm_{ik}^2} \times \sqrt{\sum_{k=1}^{n} dtm_{jk}^2}} \tag{5}$$

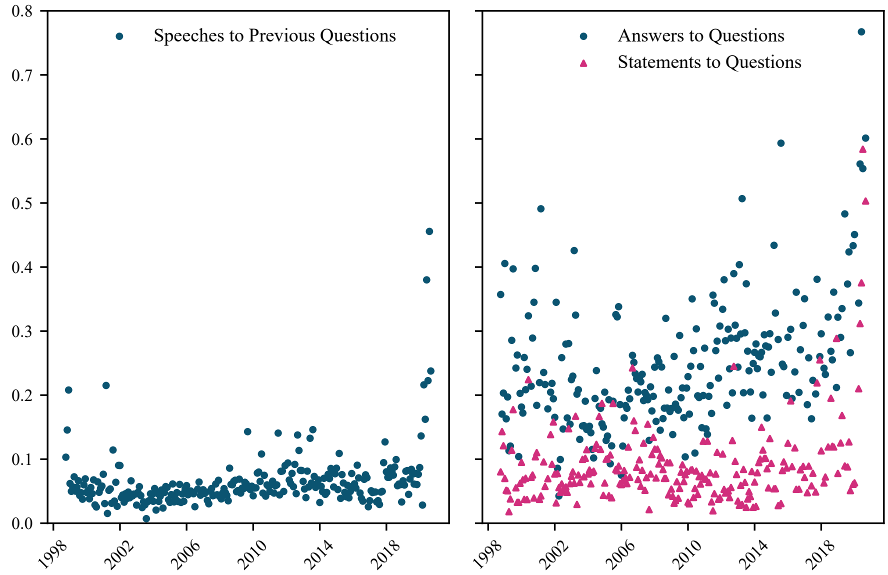
where n is the total number of terms in the document term matrix, and dtm_{ik} refers to the kth term in the vector corresponding to document i.

Figure 7, left panel, plots the similarity for speeches compared to the questions from the previous press conference ($Sim_{Sp,Q}$). Each speech day is an individual observation, but they are stacked on the date of the Governing Council meeting from which the questions are derived. There is a sharp increase in similarity towards the end of the sample. This is primarily driven by the onset of the Covid-19 crisis in 2020, where the topics of all discussions became polarised towards one specific topic. Other clusters of increased similarity are seen across the sample. A notable example is in late 2009, when a large number of speeches were made discussing the path forward for ECB monetary policy in the context of the crisis measures implemented during the financial crisis. In addition, it is noteworthy that speeches that contain references to "hearing" and "European Parliament" in their titles have, on average, a much larger recorded similarity score (0.102 compared to 0.062). This highlights the extent to which such meetings can be used to fill any information deficit.

Our argument is that Governing Council members' speeches that address the information deficit should contain more newsworthy information. Of course, ECB officials can address journalists' questions in the press conference through their direct answers at that point, so we will also assess the extent to which they do this. The fact that they had answered a question on a prevalent topic would not necessarily mean an exhaustion of all interest in that issue, particularly if it were complicated or multi-faceted, however. The right panel of Figure 7 shows the similarity scores for statements ($Sim_{S,Q}$) and answers

[24]Prior to calculation, cleaning steps are taken in line with the steps for the topic model, excluding the creation of the set of n-grams. In addition, due to the large sparse nature of the weighted document term matrix, an additional step is taken to remove the most rare and most frequent terms (those that occur in less than 1% of the documents and greater than 50% of the documents respectively). All speeches that occur on the same day as a Governing Council press conference are dropped from the sample for the calculation of similarity indices.

Figure 7: Similarity Measures



Notes: This Figure shows similarity measures constructed according to (5). Panel (a) displays the similarity between each speech and the questions in the press conference that preceded it. Panel (b) shows the similarity between the questions and (i) the answers given and (ii) the opening statement of the same press conference. For these time-series plots, we display similarity measures averaged across months.

($Sim_{A,Q}$) when both are compared to the questions in the associated press conference. Three key findings emerge. First, questions are, in general, not that similar to the statement which fits with our assumption that journalists seek missing information with their questions. Second, questions are consistently more similar to answers than to other measures. This is reassuring since the answers should, in theory, directly address the topics in the questions. Finally, a pandemic effect is present in 2020 too.

Table 8 presents the results of a regression to test whether asset price news, associated with a speech, depends on whether the speech addresses the information deficit measured using our press-questions' similarity. We include the similarity of the press conference answers to the press conference questions to capture the already-provided information, as well as its interaction with $Sim_{Sp,Q}$. The results support our hypothesis at the 6-month to 2-year maturity range; speeches addressing the information deficit give rise to more market news. However, the extent of this news may be reduced when speeches are similar to the answers given in the press conference.[25] At the 1 year maturity, a speech with similarity to the questions at the 75th percentile ($Sim_{Sp,Q} = 0.08$) generates 0.26 basis points of market reaction at an average value of answer-question similarity

[25] Table A.3 shows that these results are robust to also controlling for statement-to-questions similarity.

($Sim_{A,Q} = 0.24$). If the speech were to follow a press conference in which the answers had not addressed the questions at all ($Sim_{A,Q} = 0$), so the information deficit was larger, the market reaction would rise to approximately 1 basis point. Given that the average value of market news on ECB speech days in our sample is 2.11bp, the potential contribution of targeting communication to filling information deficits is reasonably large.

Table 8: Speech-Question similarity and Information Deficit

	(1) OIS 1M	(2) OIS 6M	(3) OIS 1Y	(4) OIS 2Y	(5) OIS 3Y	(6) DE 5Y	(7) DE 10Y
Sim. Sp. to Q.	-0.75	9.15**	12.30**	11.93*	12.48	7.81	5.90
	(5.00)	(3.90)	(5.16)	(7.01)	(7.77)	(9.11)	(8.84)
Sim. A. to Q.	-0.17	1.86	1.57	2.40	0.69	2.21	2.42
	(1.49)	(1.16)	(1.54)	(2.09)	(2.38)	(2.72)	(2.64)
Sim. Sp. to Q. × Sim. A. to Q.	-8.04	-32.19**	-37.69**	-37.90	-36.42	-28.51	-20.79
	(17.25)	(13.45)	(17.79)	(24.21)	(26.60)	(31.46)	(30.49)
Constant	1.97	5.33*	9.72**	8.92	6.62	7.38	0.25
	(4.12)	(3.21)	(4.25)	(5.78)	(6.49)	(7.49)	(7.26)
Speaker FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Topics, Future and Past	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,120	1,120	1,120	1,120	1,007	1,122	1,122
R^2	0.258	0.373	0.409	0.390	0.342	0.222	0.148
Adj. R^2	0.185	0.310	0.351	0.330	0.274	0.145	0.064

Notes: This table shows OLS estimates from regressions of the absolute value of the two day change in yields around a speech on measures of the speech similarity with the press conference questions and the Information Deficit. "$Sim_{Sp,Q}$" and "$Sim_{A,Q}$" denote the similarity of a speech and of the press conference answers to the questions, respectively. Fixed effects are included for the Executive Board member, for the year and for the day of the week in which the speech was given. Six macroeconomic surprises are included, five for the euro area and one for the US Parentheses below point estimates indicate standard errors. The statistical significance level is displayed as *p<0.1; **p<0.05; ***p<0.01.

5.2 The temporal nature of journalists' questions

Having shown that the questions are reflective of the state of the information deficit at the time, we now look at the temporal nature of this novel measure of the information deficit. Alongside our algorithm measure for topic and temporality, these questions have also been hand-coded both in terms of topic and temporal orientation.

The key summary statistics are provided below in Table 9. Over the sample, we have

6,175 total questions, of which approximately a third contain information about the past and just under two-thirds contain future information. These shares are similar for the first three sets of questions posed to the President.

Table 9: ECB Press Conference: Journalists' Questions

	Total	Past	Mixture	Future
Number	6,175	1,684	626	3,865
Percentage		27%	10%	63%
Percentage 1st 3 Q's		28%	10%	61%

Notes: This table shows summary statistics about the temporal nature of journalists' questions during the ECB press conference.

Journalists frequently ask questions on past data to enquire about (re)assessment of the state and how reassessments affect the ECB's preferences. A common example of a question about how recent data affect preferences relates observations of exchange rate movements to a possible monetary policy response. For instance, in October 1998, a journalist noted that President Duisenberg had previously said that "a fall in the dollar could be seen to have an impact [on competitiveness]" and asked "[n]ow the dollar has fallen even further and there are forecasts this will continue. What does that mean [...] for the monetary strategy of the ECB?".

Journalists frequently ask about the assessment of the state to reduce uncertainty and better communicate the ECB's view to the public. For example, in October 2007, a journalist noted to President Trichet that the recent level of inflation expectations were "higher than they been [since] the time of the euro changeover". He was then asked to clarify how these data affect his assessment and outlook. Another example comes from President Draghi's tenure, in April 2014, in which one of the first questions at the press conference sought to clarify the assessment of the state of the inflation process. The journalist noted that if one looked at producer prices rather than consumer prices, "[they were] already in a deflationary scenario", and asked how these data sources were weighed in Draghi's assessment.

A more recent example comes from October 2021, in which President Lagarde was asked whether there had been "at least a slightly different assessment to the nature of inflation given that inflation is at a 30-year high now for the [euro area]". In this instance, the journalist asked whether the policymaker might have taken a new view on how the already-observed past data might have been generated. We display the full text of these examples in Table A.4, along with three examples of reassessment questions from Federal Reserve press conferences.

6 Conclusion

Over recent decades, central banks have increased the depth, range and frequency of their communication with the public. Central bankers give speeches to the public, respond to queries from the media, interpret recent economic developments, project their intended policy path, and provide forecasts of key economic variables. A number of studies have shown that central bank communication can measurably surprise, or generate news. For communication to be effective in this way, it must, at least in part, fill some information deficit on the part of the public. The exact nature of the information deficit has been less studied in the literature.

Some earlier studies supposed that the information deficit was related to private information on the part of the central bank. However, at least in the narrow interpretation of private information, markets have broadly the same access to data that the central bank has, reducing the scope for substantively new information of this type to be released by the central bank. We argue instead that the source of the information advantage that generates the news in central bank communication comes from the central bank updating either its evaluation of the current state of the economy, or its mapping of the state into the appropriate monetary policy stance.

Our analysis of the temporal dimension of communication, facilitated by our new measurement methodology, suggests that the information communicated by central banks that is relevant for forward-looking expectations is multi-faceted and is not reducible to the numerical forecast data that they also publish, or to explicit policy forward guidance. While central banks may have no advantage purely in terms of methodology or training than markets, they often have an advantage in terms of the resources: the number of staff and models available to generate the assessment of the state. Hence the central bank can devote resources to *evaluation* of economic data - choosing the weights to put on different data sources at each point in time and how to place developments in the data into context. This process is inherently backward-looking.

An implication of having an important role for an evolving (time-varying) assessment is that the market could always be in deficit. This means that the communication of a fixed reaction function, which is highly desirable in many FIRE models, is likely impossible in practice. This is because of time-variation in the way that the macroeconomic data maps to the likely state of the world to which policy has to respond. This is true even if preferences remain fixed over time, but preferences may also evolve. As emphasised by Williams (2019) when thinking about the data-dependence of monetary policy: "I wish I could now tell you with certainty what will happen to the economy, but anyone who promises they can see into the future is a charlatan. However, what I can do is provide you with some insight into how I assess the health of the economy and what that means for my view on the monetary policy decisions before us."

While this means that central banks *could* always provide useful information, it does not follow that the central bank *should* provide constant updates to the markets. It is likely that the central bank is also learning from the data and exploring possible narratives that explain the observed data. In this environment, and given this process is subject to noise, it might be optimal only to incrementally update the market with their views. We leave this question of how often to communicate updated views to future research.

But our work does indicate what is important to communicate. Even though markets may be solely interested in the outlook for future monetary policy, their interest will be more than just forward guidance. Policymakers should try to communicate how they are assessing incoming data, and how this affects their thinking about both the current (evaluation) and future (projection) state of the economy. This goes beyond the recommendations from standard models used in the analysis of monetary policy.

Bibliography

Aguilar, A. and F. Pérez-Cervantes (2022, June). Communication, monetary policy, and financial markets in Mexico. BIS Working Papers 1025, Bank for International Settlements.

Altavilla, C., L. Brugnolini, R. S. Gürkaynak, R. Motto, and G. Ragusa (2019). Measuring euro area monetary policy. *Journal of Monetary Economics 108*, 162–179.

Andre, P., C. Pizzinelli, C. Roth, and J. Wohlfart (2022, 02). Subjective Models of the Macroeconomy: Evidence From Experts and Representative Samples. *The Review of Economic Studies 89*(6), 2958–2991.

Angelico, C., J. Marcucci, M. Miccoli, and F. Quarta (2022). Can we measure inflation expectations using twitter? *Journal of Econometrics 228*(2), 259–277.

Apel, M., M. Blix Grimaldi, and I. Hull (2019, November). How Much Information Do Monetary Policy Committees Disclose? Evidence from the FOMC's Minutes and Transcripts. Working Paper Series 381, Sveriges Riksbank (Central Bank of Sweden).

Arismendi-Zambrano, J., E. Kypraios, and A. Paccagnini (2021, November). Federal Reserve Chair Communication Sentiments' Heterogeneity, Personal Characteristics, and their Impact on Uncertainty and Target Rate Discovery. Technical Report ICM-2021-01, Henley Business School, ICMA Centre.

Baker, S. R., N. Bloom, and S. J. Davis (2016, November). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Bauer, M. D., A. K. Lakdawala, and P. Mueller (2021, November). Market-Based Monetary Policy Uncertainty. *The Economic Journal*.

Bauer, M. D. and E. T. Swanson (2020, April). An alternative explanation for the "fed information effect". Working Paper 27013, National Bureau of Economic Research.

Bauer, M. D. and E. T. Swanson (2022, May). A reassessment of monetary policy surprises and high-frequency identification. Technical report, mimeo.

Bennani, H. and M. Neuenkirch (2017). The (home) bias of european central bankers: new evidence based on speeches. *Applied Economics 49*(11), 1114–1131.

Björkelund, A., B. Bohnet, L. Hafdell, and P. Nugues (2010). A High-Performance Syntactic and Semantic Dependency Parser. Technical report, Coling 2010: Demonstrations.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003, March). Latent dirichlet allocation. *Journal of Machine Learning Research 3*, 993–1022.

Boukus, E. and J. V. Rosenberg (2006, January). The information content of fomc minutes. Technical report, Federal Reserve Bank of New York.

Bundick, B., T. Herriford, and A. L. Smith (2021, February). Forward Guidance, Monetary Policy Uncertainty, and the Term Premium. Research Working Paper 17-07, Federal Reserve Bank of Kansas City.

Bybee, L., B. Kelly, A. Manela, and D. Xiu (2020, January). The structure of economic news. *mimeo.*

Byrne, D., R. Goodhead, M. McMahon, and C. Parle (2023, February). Measuring the Temporal Dimension of Text: An Application to Policymaker Speeches. Research Technical Paper 2023/02, Central Bank of Ireland.

Caldara, D. and M. Iacoviello (2022, April). Measuring geopolitical risk. *American Economic Review 112*(4), 1194–1225.

Cannon, S. A. (2015). Sentiment of the FOMC: Unscripted. *Economic Review* (Q IV), 5–31.

Carney, M. (2017, 1). Lambda. Speech given by Mark Carney, Governor of the Bank of England.

Chang, A. X. and C. Manning (2012, May). SUTime: A library for recognizing and normalizing time expressions. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, pp. 3735–3740. European Language Resources Association (ELRA).

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

Clarida, R., J. Galí, and M. Gertler (1999, December). The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature 37*(4), 1661–1707.

Coenen, G., M. Ehrmann, G. Gaballo, P. Hoffmann, A. Nakov, S. Nardelli, E. Persson, and G. Strasser (2017, June). Communication of monetary policy in unconventional times. Working Paper Series 2080, European Central Bank.

Cross, J. P. and D. Greene (2020). Talk is not cheap: Policy agendas, information processing, and the unusually proportional nature of european central bank communications policy responses. *Governance 33*, 425–444.

Das, S. R. and M. Y. Chen (2007, September). Yahoo! for amazon: Sentiment extraction from small talk on the web. *Management Science 53*(9), 1375–1388.

De Pooter, M., G. Favara, M. Modugno, and J. Wu (2021). Monetary policy uncertainty and monetary policy surprises. *Journal of International Money and Finance 114*(102401).

Draghi, M. (2019). Introductory statement, press conference, 12 september 2019. https://www.ecb.europa.eu/press/pressconf/2019/html/ecb.is190912~658eb51d68.en.html.

Eggertsson, G. B. and M. Woodford (2003). The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity 2003*(1), 139–211.

Ehrmann, M. and M. Fratzscher (2009, June). Explaining Monetary Policy in Press Conferences. *International Journal of Central Banking 5*(2), 42–84.

Ehrmann, M. and D. Sondermann (2012, September). The News Content of Macroeconomic Announcements: What if Central Bank Communication Becomes Stale? *International Journal of Central Banking 8*(3), 1–53.

Ellingsen, T. and U. Söderström (2001, December). Monetary policy and market interest rates. *American Economic Review 91*(5), 1594–1607.

Flynn, J. P. and K. Sastry (2022). The macroeconomics of narratives. *Available at SSRN 4140751*.

Galardo, M. and C. Guerrieri (2017, July). The effects of central bank's verbal guidance: evidence from the ECB. Temi di discussione (Economic working papers) 1129, Bank of Italy, Economic Research and International Relations Area.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton, USA and Oxford, UK: Princeton University Press.

Gentzkow, M., B. Kelly, and M. Taddy (2019, September). Text as Data. *Journal of Economic Literature 57*(3), 535–574.

Gertler, M. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Goetzmann, W. N., D. Kim, and R. J. Shiller (2022, July). Crash narratives. Working Paper 30195, National Bureau of Economic Research.

Gorodnichenko, Y., T. Pham, and O. Talavera (2021, March). The Voice of Monetary Policy. NBER Working Papers 28592, National Bureau of Economic Research, Inc.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005, May). Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. *International Journal of Central Banking 1*(1).

Hansen, S. and M. McMahon (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics 99*(S1), S114–S133.

Hansen, S., M. McMahon, and A. Prat (2017). Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics*. Forthcoming.

Hansen, S., M. McMahon, and M. Tong (2019). The long-run information effect of central bank communication. *Journal of Monetary Economics 108*, 185–202.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2019, November). Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics 134*(4), 2135–2202.

Hayo, B. and M. Neuenkirch (2010). Do federal reserve communications help predict federal funds target rate decisions? *Journal of Macroeconomics 32*, 1014–1024.

Hendry, S. and A. Madeley (2010, November). Text mining and the information content of bank of canada communications. Technical Report 31, Bank of Canada.

Hubert, P. and F. Labondance (2021). The signaling effects of central bank tone. *European Economic Review 133*(C).

Istrefi, K., F. Odendahl, and G. Sestieri (2021). Fed communication on financial stability concerns and monetary policy decisions: revelations from speeches. Working Papers 2110, Banco de España.

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

Kocherlakota, N. (2016). Rules versus discretion: A reconsideration. *Brookings Papers on Economic Activity*, 1–36.

Kuttner, K. N. (2001, June). Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics 47*(3), 523–544.

Larsen, V. H., L. A. Thorsrud, and J. Zhulanova (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics 117*, 507–520.

Loughran, T. and B. McDonald (2011, February). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance 66*(1), 35–65.

Lucca, D. O. and F. Trebbi (2009, September). Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. NBER Working Papers 15367, National Bureau of Economic Research, Inc.

Macaulay, A. and W. Song (2022). Narrative-driven fluctuations in sentiment: Evidence linking traditional and social media. *Available at SSRN 4150087*.

McMahon, M. and T. Munday (2022a). The element(s) of surprise. Working paper, Mimeograph.

McMahon, M. and T. Munday (2022b). Uncertainty and time-varying monetary policy. Working paper, Mimeograph.

Nakamura, E. and J. Steinsson (2017). High frequency identification of monetary non-neutrality: The information effect. *The Quarterly Journal of Economics* (Forthcoming).

Orphanides, A. (2001). Monetary policy rules based on real-time data. *The American Economic Review 91*(4), 964–985.

Orphanides, A. and J. C. Williams (2005). Inflation scares and forecast-based monetary policy. *Review of Economic Dynamics 8*(2), 498–527.

Parle, C. (2022). The financial market impact of ecb monetary policy press conferences — a text based approach. *European Journal of Political Economy 74*, 102230.

Picault, M. and T. Renault (2017). Words are not all created equal: A new measure of ecb communication. *Journal of International Money and Finance 79*, 136–156.

Ramm, A., S. Loáiciga, A. Friedrich, and A. Fraser (2017, July). Annotating tense, mood and voice for English, French and German. In *Proceedings of ACL 2017, System Demonstrations*, Vancouver, Canada, pp. 1–6. Association for Computational Linguistics.

Romer, C. D. and D. H. Romer (2000, June). Federal reserve information and the behavior of interest rates. *American Economic Review 90*(3), 429–457.

Rosa, C. and G. Verga (2007). On the consistency and effectiveness of central bank communication: Evidence from the ecb. *European Journal of Political Economy 23*(1), 146–175. Central bank transparency and central bank communication.

Schmeling, M. and C. Wagner (2019, January). Does Central Bank Tone Move Asset Prices? CEPR Discussion Papers 13490, C.E.P.R. Discussion Papers.

Shapiro, A. H., M. Sudhof, and D. J. Wilson (2020, March). Measuring News Sentiment. Working Paper Series 2017-1, Federal Reserve Bank of San Francisco.

Shiller, R. J. (2017, April). Narrative Economics. *American Economic Review 107*(4), 967–1004.

Sims, C. A. and T. Zha (2006, March). Were There Regime Switches in U.S. Monetary Policy? *American Economic Review 96*(1), 54–81.

Taddy, M. (2017). One-Step Estimator Paths for Concave Regularization. *Journal of Computational and Graphical Statistics 26*(3), 525–536.

Tetlock, P. (2007, June). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance 62*, 1139–1168.

Tobback, E., S. Nardelli, and D. Martens (2017). Between hawks and doves: measuring central bank communication. Working Paper Series 2085, European Central Bank.

Williams, J. (2019). Monetary policy—a "data dependent" approach. At the New Jersey Bankers Association's Economic Leadership Forum, Somerset, New Jersey, 18 January 2019.

Zou, H. and T. Hastie (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society 67*, 301–320. Part 2.

A Appendix

A.1 Estimated Topics

Table A.1: Topic Model: Federal Reserve

1 Labour	2 Inflation and Risk	3 Inflation	4 Policy Rates	5 Procedural
labor	measur	price	rate	bank
market	inflat	energi	feder	feder_reserv
rate	project	inflat	fund	messr
remain	risk	increas	committe	action
inflat	particip	import	percent	presid
unemploy	rang	declin	target	economist
particip	uncertainti	consum	rang	vote
pressur	year	month	level	respect
recent	inflat_expect	percent	inflat	general
increas	forecast	core	market	counsel

6 Open Economy	7 Committee I	8 Growth	9 Securities Markets	10 Forecasts
market	board	growth	secur	rate
foreign	governor	economi	agenc	project
oper	divis	busi	loan	year
system	director	econom	credit	unemploy
open	monetari	particip	treasuri	percent
currenc	affair	activ	hold	particip
account	secretari	spend	purchas	staff
bank	assist	continu	mortgage-back	forecast
manag	research	effect	continu	estim
committe	offic	member	bank	longer-run

11 Monetary Policy	12 Financial Markets	13 Committee II	14 Policy Outlook	15 Real Activity Data
polici	period	committe	committe	quarter
particip	market	market	econom	busi
well	intermeet	meet	inflat	spend
economi	yield	open	polici	sale
monetari_polici	dollar	feder	outlook	increas
that	declin	vote	member	pace
use	spread	direct	continu	continu
provid	price	unanim	inform	product
effect	equiti	releas	market	month
chang	treasuri	held	monetari_polici	declin

A.2 The Relation Between Monetary and Information Surprises and Temporal Information

Figure A.2 plots for the ECB data the full distributions of Adjusted R-Squared values from Post-LASSO Ordinary Least Squares regressions that underpin Table **??**. The left column shows the regressions for the exercise in which only topics without any temporal dimension are included in the topic set ("Topics Only" row of the table), while the right column show the most disaggregate results with both future and past temporal topics ("Topics, Future and Past*" row).[26] The top row, Figures A.2a and A.2b, show the results for the information surprise, "INFO", and each figure includes the findings from regressing on the set of variables selected by LASSO for the information surprise itself and the set selected for the monetary policy surprise ("MPOL") for the same bootstrap draw. Figures A.2c and A.2d repeat these results but where the monetary policy surprise ("MPOL") is the dependent variable.

Two main results stand out. The first is that the topics, while capturing broad themes, are too broad to capture a lot of the market news. This is reflected in both the fact the Adjusted R^2 is relatively low at around 0.2 (as in the Table). The second is that there is little extra explanatory power from using the topics selected by LASSO for one surprise in explaning the other surprise – the distributions are very similar. When we instead use the disaggregated temporal shocks, the Adjusted R^2 jumps up to around 0.6-0.7 for the own shocks, and the specificity of the selected is much greater – using the topics selected for the other shock, almost halves the explanatory power. This shows that the communication that moves markets is best captured in high-dimensional measures of the messaging rather than simply broad themes.

A.3 ECB Speeches: Baseline Analysis

As described in the main text, we repeat the baseline analysis to the speeches. When studying the effect of the speeches on financial market variables, we use daily (end-of-day) series, downloaded from Bloomberg. There are several reasons for this. The first is that many of the speeches are given outside market trading hours, meaning the construction of an intra-daily movement is impossible. The second is that we do not know exactly when the information contained within the speeches became generally available to markets, since this information is not recorded in the dataset.[27] For our empirical specifications, we drop 90 speech-day observations that fall on Saturdays or Sundays, and we also drop

[26] Forecast control variables are constrained to be always included in the set of selected variables.

[27] Note that for a speech delivered on a Friday, we employ a window from market-close on Thursday to market-close on Monday. We account for potential heterogeneity in the treatment of speeches across days using day-of-the-week fixed effects in our empirical specifications.

Table A.2: Topic Model: ECB

1	2	3	4	5
Risk	Structural Reform	Prices and Inflation	Banks, Markets, Bonds	EA Growth
risk	euro_area	price	bank	euro_area
uncertainti	structur_reform	increas	measur	continu
econom	market	inflat	market	support
euro_area	competit	effect	bond	growth
outlook	economi	expect	monetari_polici	expect
relat	increas	year	programm	demand
financi_market	countri	month	certain	loan
develop	need	oil_price	differ	remain
high	growth	inflat_rate	liquid	recoveri
growth	product	energi	effect	credit

6	7	8	9	10
Price Stability	Price Stability and Expectations	Gov. Co. Logistics	Discursive	Data
price_stabil	price_stabil	meet	well	year
monetari_polici	inflat_expect	govern_council	go	quarter
govern_council	medium_term	ecb	chang	growth
risk	close	confer	comment	euro_area
monitor	line	presid	thing	data
develop	inflat	outcom	last	indic
decis	deliv	introductori	certain	last
close	mandat	decis	ask	confirm
stanc	maintain	staff	actual	assess
assess	continu	project	discuss	inform

11	12	13	14	15
Monitoring	Fiscal	ECB	Forward Guidance	Rate Decisions
import	countri	ecb	rate	rate
situat	govern	central_bank	continu	interest_rate
observ	fiscal	govern_council	remain	decid
market	growth	european	monetari_analysi	key
regard	stabil	euro	growth	basi
cours	euro_area	member	under	ecb_interest_rate
economi	fiscal_polici	institut	low	oper
europ	implement	presid	monetari	point
level	pact	decis	money	loan
present	need	govern	confirm	accord

Table A.3: Speech-Question similarity and Information Deficit - Robustness to Statements

	(1) OIS 1M	(2) OIS 6M	(3) OIS 1Y	(4) OIS 2Y	(5) OIS 3Y	(6) DE 5Y	(7) DE 10Y
Sim. Sp. to Q.	-1.56	8.58**	11.64**	11.34	11.74	7.48	7.16
	(5.11)	(3.98)	(5.26)	(7.15)	(7.91)	(9.28)	(9.00)
Sim. A. to Q.	0.21	2.29*	2.04	3.05	1.33	3.13	2.72
	(1.55)	(1.21)	(1.60)	(2.17)	(2.47)	(2.82)	(2.73)
Sim. S. to Q.	-2.14	-3.30*	-3.53	-5.61*	-4.80	-9.17**	-6.42
	(2.40)	(1.87)	(2.48)	(3.37)	(3.67)	(4.37)	(4.23)
Sim. Sp. to Q. × Sim. A. to Q.	-12.75	-35.45**	-41.52**	-41.34	-41.04	-30.46	-13.55
	(18.26)	(14.22)	(18.82)	(25.59)	(28.07)	(33.20)	(32.18)
Sim. Sp. to Q. × Sim. S. to Q.	20.91	16.80	19.28	20.41	23.40	19.14	-19.84
	(25.07)	(19.52)	(25.84)	(35.13)	(37.31)	(45.58)	(44.17)
Constant	2.24	5.57*	9.99**	9.24	7.24	7.70	0.11
	(4.13)	(3.22)	(4.26)	(5.79)	(6.52)	(7.49)	(7.26)
Speaker FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Topics, Future and Past	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,120	1,120	1,120	1,120	1,007	1,122	1,122
R^2	0.259	0.375	0.411	0.392	0.344	0.227	0.154
Adj. R^2	0.184	0.311	0.351	0.331	0.274	0.149	0.068

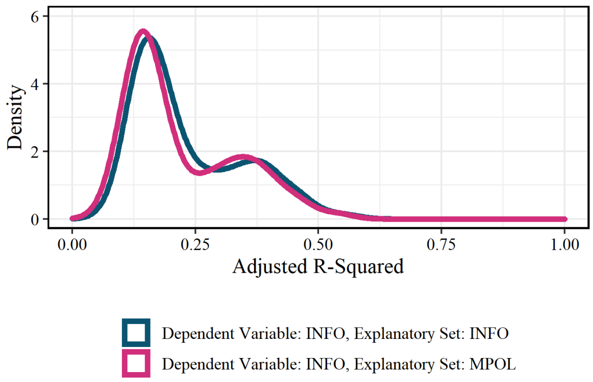
Notes: This table shows OLS estimates from regressions of the absolute value of the two day change in yields around a speech on measures of the speech similarity with the press conference questions and the Information Deficit. "$Sim_{Sp,Q}$", "$Sim_{A,Q}$" and "$Sim_{S,Q}$" denote the similarity of a speech, of the press conference answers and of the Introductory Statement to the questions, respectively. Fixed effects are included for the Executive Board member, for the year and for the day of the week in which the speech was given. Six macroeconomic surprises are included, five for the euro area and one for the US Parentheses below point estimates indicate standard errors. The statistical significance level is displayed as *p<0.1; **p<0.05; ***p<0.01.

Table A.4: Examples of Journalist Questions Regarding Central Bank Evaluation

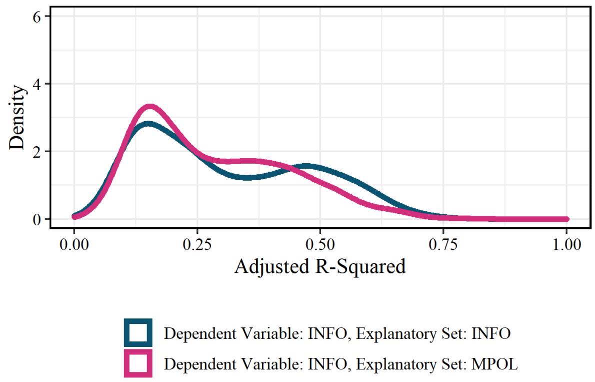
Central Bank	Date	Question
ECB	04/10/2007	You mentioned inflation expectations as something very important to the Governing Council. If I have got this right, inflation expectations in a number of countries are actually quite high, higher than they have been or at levels around at the time of the euro changeover. Can you talk a little bit about how that is figuring into your expectations and into your calculations?
ECB	03/04/2014	A first question relating to the inflation rate and referring not so much to consumer prices but to producer prices, where we have seen two consecutive months of a decline – so, if you take producer prices, you could actually argue that we are already in a deflationary scenario. To what extent would you consider producer prices and how important are they for your assessment?
ECB	28/10/2021	I was wondering, because the markets were – well, everybody was anticipating that this meeting was all about inflation, so what have you been discussing, and what was on top of your agenda? Was there at least a slightly different assessment to the nature of inflation, given that inflation is at a 30-year high now for the eurozone.
Federal Reserve	25/01/2012	Mr. Chairman, we've had several months of economic data that's been stronger than most forecasters expected–employment was over 200,000, the unemployment rate's come down to 8.5 percent–but there seems to be very little mention of this recent strength in the statement. Do you and the Committee, Mr. Chairman, harbor doubts about the recent strength in the economy?
Federal Reserve	19/03/2014	You mentioned in your testimony on Capitol Hill recently that the Fed was trying to assess the balance of weather effects versus more fundamental weakness in the economy as the reason for the slowdown in growth in the first quarter, and you guys mentioned in the statement weather specifically. Does that mean that the Fed's analysis has come down on the side of weather, or are you still concerned that there could be something else going on that could be contributing to slower growth?
Federal Reserve	30/01/2019	I am struggling a little bit to understand what has changed since we sat here with you six weeks ago. You've said today that you think that inflation would be the reason that the Fed would need to continue raising rates. Has the inflation outlook shifted that dramatically in the last six weeks?

Figure A.1: Identifying Distinct Information Sets Using Topics and Temporal Topics Federal Reserve

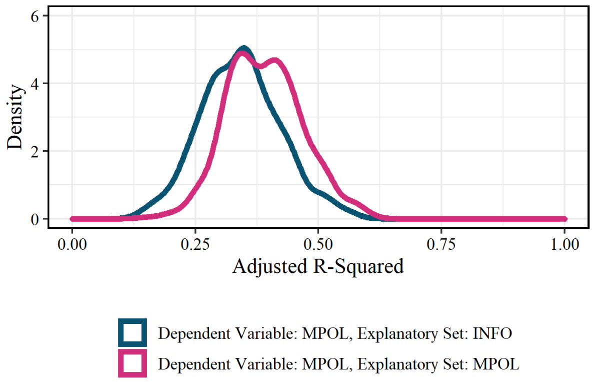
(a) INFO: Topics Only



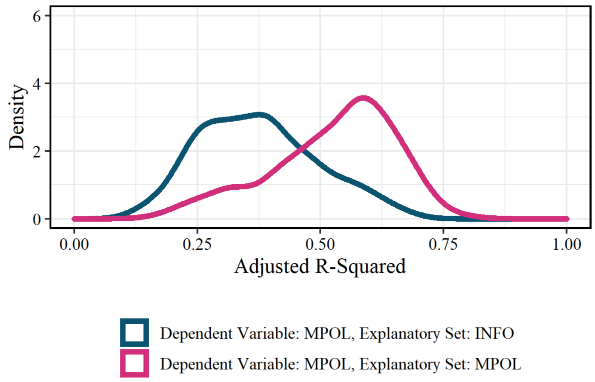
(b) INFO: Topics, Future and Past



(c) MPOL: Topics Only

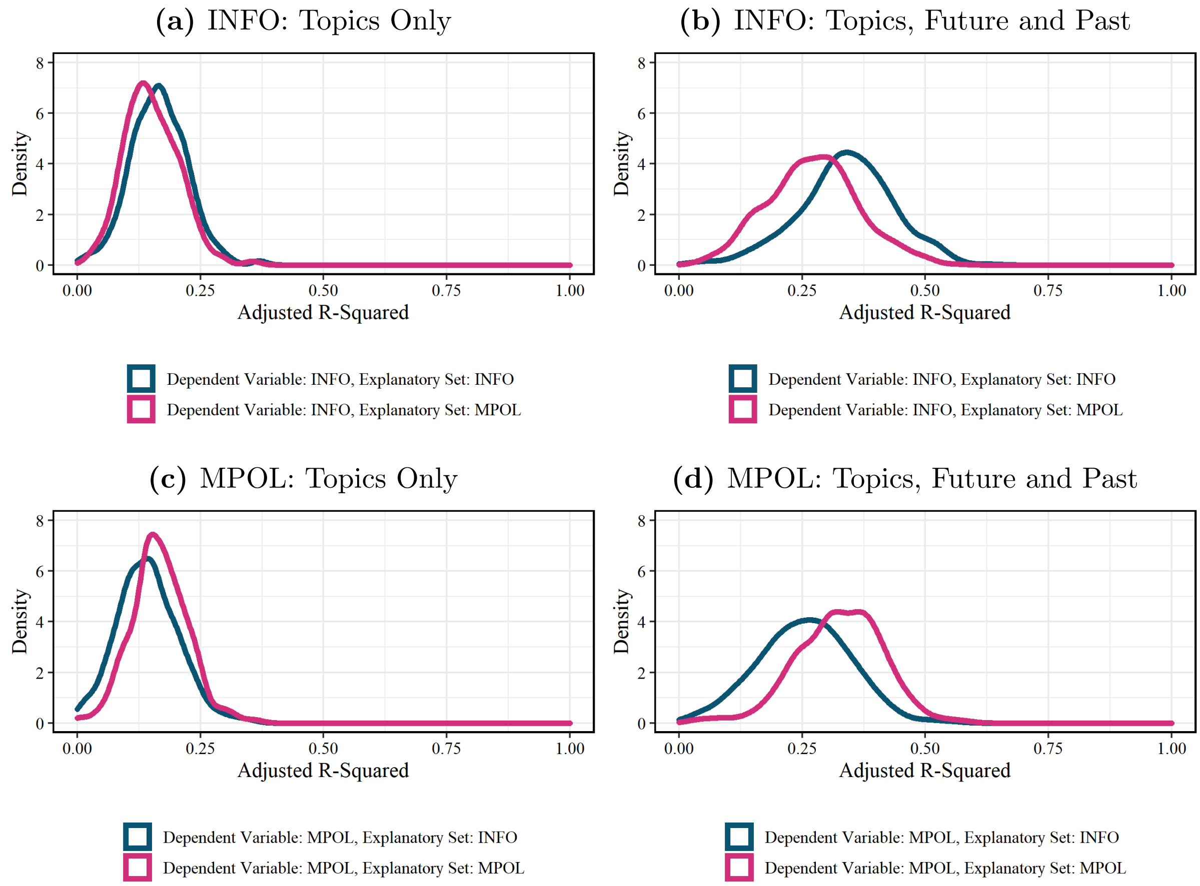


(d) MPOL: Topics, Future and Past



Notes: The top row shows the distributions of Adjusted R-Squared values from Post-LASSO Ordinary Least Squares regressions of the information surprise ("INFO") on the set of variables selected by LASSO for the information surprise itself and the set selected for the monetary policy surprise ("MPOL") for the same bootstrap draw. The set of variables available to be selected includes the Topics only (θ_k). The bottom row shows the distributions of Adjusted R-Squared across the bootstrap draws when the set of variables available to be selected has Topics (θ_k), Future (θ_k^{FUT}) and Past (θ_k^{PAST}). Each specification includes forecasts of current year and one year ahead annual GDP growth and inflation, and revisions to these forecasts where applicable. The forecast variables are constrained to be always included in the set of selected variables.

Figure A.2: Identifying Distinct Information Sets Using Topics and Temporal Topics ECB



Notes: The top row shows the distributions of Adjusted R-Squared values from Post-LASSO Ordinary Least Squares regressions of the information surprise ("INFO") on the set of variables selected by LASSO for the information surprise itself and the set selected for the monetary policy surprise ("MPOL") for the same bootstrap draw. The set of variables available to be selected includes the Topics only (θ_k). The bottom row shows the distributions of Adjusted R-Squared across the bootstrap draws when the set of variables available to be selected has Topics (θ_k), Future (θ_k^{FUT}) and Past (θ_k^{PAST}). Each specification includes forecasts of current year and one year ahead annual GDP growth and inflation, and revisions to these forecasts where applicable. The forecast variables are constrained to be always included in the set of selected variables.

150 speech-day observations that fall on Governing Council meeting days, the day before, or the day after. This means the total number of speech-day observations available for analysis is 1,473.

Table A.5b presents the main analysis; it repeats the baseline analysis of Table **??** but applied to the speeches. The temporal dimension of speeches generates similar, albeit less systematic, market news.

Table A.5: Policymaker Speeches

(a) Federal Reserve

Specification	ED1	ED4	US 2Y	US 5Y	US 10Y
Controls Only	0.245	0.237	0.176	0.209	0.173
Topics	0.261	0.255	0.199	0.228	0.194
Topics, Future Topics	0.265	0.259	0.203	0.232	0.198
Topics, Past Topics	0.264	0.256	0.201	0.229	0.195
Topics, Future and Past Topics	0.268	0.260	0.205	0.233	0.200
Topics, Future and Past Topics*	0.321	0.314	0.262	0.278	0.249

(b) ECB

Specification	OIS 1M	OIS 1Y	OIS 2Y	DE 5Y	DE 10Y
Controls Only	0.198	0.329	0.322	0.255	0.146
Topics	0.243	0.367	0.360	0.292	0.186
Topics, Future Topics	0.251	0.370	0.363	0.297	0.197
Topics, Past Topics	0.254	0.371	0.365	0.298	0.195
Topics, Future and Past Topics	0.262	0.375	0.369	0.304	0.205
Topics, Future and Past Topics*	0.289	0.411	0.411	0.347	0.252

Notes: This table shows the adjusted R^2 for specifications predicting the responses of the absolute value of two-day changes in yields at different maturities, respectively for the Federal Reserve (top panel) and the ECB (lower panel). The events studied in this exercise relate to policymaker speech days. Specifications are identical to Table 3, excepting the changes in events, the dependent variables, and differing control variables. Daily financial data was sourced from Bloomberg. See Notes to Table 3 for further details regarding the Elastic Net specifications. The additional specification, "Topics, Future and Past Topics*" refers to a case where we include controls, topics, and temporal topics. In this case the temporal topics are disaggregated, meaning we have temporal topics based separately on tags from categorical, numerical, and grammatical tags, respectively for past and future.