The Long-Run Information Effect of Central Bank Text*

Stephen Hansen† Michael McMahon‡ Matthew Tong§

December 29, 2017

Preliminary —Please Do Not Cite

Abstract

Central bank communication is an increasingly important policy tool, but its impact on market expectations arises through many potential channels. We clarify how communicating solely about expected future economic conditions can move short- and long-run market rates, and use the publication of the Bank of England's Inflation Report to study the market impact of such communication. We find compelling evidence for a long-run information effect driven more by narrative than by quantitative forecasts. The effect operates largely via its impact on the term premium. We conclude that central banks can impact long-run interest rates without making explicit future policy commitments.

Keywords: Monetary Policy, Communication, Machine Learning

JEL Codes: E52, E58, C55

*We would like to thank, for comments, discussion and suggestions, Michael Ehrmann and Jonathan Talmi as well as seminar participants at the Banco do Brazil, Bank of England, ECB, Empirical Macroeconomics Conference at Ghent University, Higher School of Economics, New Economic School, Norges Bank, Oxford/NY Fed Conference, RCEA conference, and UPenn. This views expressed here do not necessarily represent the views of the Bank of England, the Monetary Policy Committee, Financial Policy Committee or anyone else other than the authors. Any errors remain ours alone.

†University of Oxford, Alan Turing Institute and CEPR. Email: stephen.hansen@economics.ox.ac.uk

‡University of Oxford, CEPR, CAGE (Warwick), CfM (LSE), and CAMA (ANU). Email: michael.mcmahon@economics.ox.ac.uk

§Bank of England. Email: matthew.tong@bankofengland.co.uk

1 Introduction

Central banks in the major economies communicate frequently nowadays, but this is a relatively recent phenomenon. As recently as January 1994, the Federal Open Market Committee (FOMC) didn't immediately announce their monetary policy decisions. The main reason appears to be a realization among central banks that expectations are a key link between their policy instruments and the interest rates that really matter for economic activity, and that communication could aid expectations management (Woodford 2001, Blinder 2008). While initially the focus was on building credibility through accountability and the adoption of specific frameworks such as inflation targeting (a form of central bank communication), more recent interest has focused on greater transparency and more frequent communication.

There is well-established evidence that central bank communication does indeed drive interest rates of different maturities (Gürkaynak et al. 2005, Boukus and Rosenberg 2006, Blinder et al. 2008, Carvalho et al. 2016, for example), but the channels through which that occurs are often unclear. The same communication event can transmit information about the current policy stance, the central bank's views on economic conditions, and the central bank's future policy stance (Campbell et al. 2012). Moreover, central banks often release several pieces of information simultaneously, for example an announcement of current policy and an accompanying statement. The extensive literature that measures monetary policy news based on high-frequency movements in short-term policy rates (Gertler and Karadi 2015, Nakamura and Steinsson 2017, Jarociński and Karadi 2017, for example) captures the aggregate impact of these separate, heterogeneous information sources. Each source can itself have multiple effects: an announcement of a policy rate decision both informs markets of the current policy stance and can signal the central bank's views of economic conditions that led to that stance (Romer and Romer 2000, Melosi 2017, Zhang 2017).

One specific example of the importance of understanding the channels through which communication operates is the ongoing debate about why long-run market interest rates react to short-run monetary news. Nakamura and Steinsson (2017) argue that this is due to monetary policy shocks transmitting information about economic fundamentals that affects long-run market expectations, termed the *information effect* by Romer and Romer (2000). In contrast, Hanson and Stein (2015) argue that news about short-term policy expectations is propagated to longer-maturity bonds by the behavioral response of yield-oriented investors. According to their model, decreases (increases) in short rates induce these investors to switch to (from) longer-maturity bonds, driving the yields on such bonds down (up) through changes in the term premium. Given the current set

of empirical methods in the literature, it is a serious challenge to precisely identify the transmission mechanism and duration of the information effect.

The first goal of this paper is to bring greater clarity about the multiplicity of effects that central bank communication can have, which we discuss within a simple illustrative framework. We emphasise the role for expectations of future policy maker's beliefs about economic conditions. Communication about current economic forecasts may, depending on persistence, directly drive expectations. But perhaps more importantly, we highlight the indirect role played by economic forecasts, and the narrative text around them, to help identify other, more persistent, drivers of future interest rates such as beliefs about real interest rates.

Our second goal is to study a controlled environment in which we isolate communication solely about economic fundamentals, so-called *Delphic* forward guidance in the language of Campbell et al. (2012). Specifically, we exploit the unique environment of the publication of the Bank of England's Inflation Report (IR) from February 1998 through May 2015. The IR contains information about the Bank of England's economic forecasts, but does not provide explicit forward guidance on future policy. Moreover, during our sample period, the IR was published according to a fixed, quarterly schedule one week *after* the announcement of the contemporaneous policy decision. It therefore constitutes a policy-free central bank information shock. This allows us to directly assess the market impact of news about the Bank of England's private views without having to decompose a policy change into separate information and policy shock components.[1]

The IR is also of interest due to its rich information structure. It contains detailed, quantitative inflation and output forecasts in the form of fan charts that place probabilities on different future realizations that are summarized with mode, variance, and skew. Much of the information effect literature focuses on markets' updating their beliefs based on such quantitative information. The IR also contains rich, narrative information in the form of text data that both describes recent economic developments and the near-term outlook, as well as the forecast itself. This structure allows us to assess the market impact of narrative information while controlling for the impact quantitative information.

To study the effect of the IR publication, we adopt an event-study methodology made popular by Cook and Hahn (1989), Kuttner (2001) and Bernanke and Kuttner (2005). We use daily data on four nominal rates at varying maturities derived from UK government bond prices: the one-year spot rate; the three-year forward rate; the five-year forward rate; and the five-year ahead, five-year forward rate. We then analyze the extent to which information in the IR can explain the absolute values of changes in these rates on the publication day, which is our measure of news. We find that the quantitative information

[1]See Miranda-Agrippino and Ricco (2015) for an example of the latter.

has strong explanatory power for explaining the one-year spot rate, but much less for longer rates. This is consistent with information on the quantitative forecasts' having an information effect at the short run, but casts doubt on whether it can substantially explain long-run rate movements.

We next ask whether narrative information can explain the residual moves in asset prices, conditional on the quantitative information published. For each unique two-word phrase in the IR, we regress its count over time on the same quantitative forecast information we use to explain asset price movements, and back out the residual. This construction is the textual analogue of the method for computing monetary policy shocks of Romer and Romer (2004), and it yields text shocks that purge from language the endogenous variation due to the current economic outlook. We then face a high-dimensional statistical problem: there are thousands of potentially relevant text characteristics and only seventy (quarterly) publication dates in our sample. To test for information content, we examine the number of significant characteristics identified by the Least Absolute Shrinkage and Selection Operator (LASSO) via cross validation and compare that to the number identified when we randomly permute the data a re-estimate the LASSO. We find that for the true data it essentially selects as many characteristics as observations, the maximum the LASSO can, but for the vast majority of randomly permuted datasets there are no text characteristics selected for any of the interest rate maturities. We therefore overwhelmingly reject the hypothesis that narrative information is unrelated to the size of the interest rate moves at both short and long-term maturities.

While a useful starting point, our baseline test is not sufficient to conclude that the narrative has a direct informational impact on long-run rates rather than simply being an information effect on short-term rates propagated to long-term rates by market mechanics. Instead, we need to show that *different* narrative information drives short- and long-term rates. To answer this question, we reduce the dimensionality of the text with Latent Dirichlet Allocation (LDA), a popular probabilistic topic model (Blei et al. 2003) previously applied in monetary policy by Hansen and McMahon (2016) and Hansen et al. (2017). LDA represents each IR as a distribution over a finite set of topics that capture common themes in the data. For example, a particular IR might devote 20% of its space to inflation; 15% of its space to financial market conditions; and so on.

We then construct topic content residuals in the same way to text residuals above, and use a bootstrapping procedure to rank topics separately for each asset maturity according to their ability to robustly predict price movements. The ranking of topics for the one-year spot rate is uncorrelated with the ranking for all longer-term rates. Moreover, the topics that explain longer-term rate movements appear to be more associated with narrative around the forecast rather than with narrative around current economic conditions. This

evidence points to a long-run information effect mediated by the text narrative separate from any effect due to quantitative forecasts.

Finally, we ask whether the long-run information effect of central bank text is driven more by changes in long-run expectations of the modal rate or the long-run term premium. We use four affine term structure models to decompose rate changes into expectation and term premium components and repeat our analysis by component. Our baseline test of information indicates that both components react to IR information at all maturities. But the narrative appears particularly important for explaining movements in the long-run term premium, with just a handful of key topics driving more variation than the more than one dozen quantitative variables. This is consistent with theories of macroeconomic conditions driving the term premium (Cochrane 2011, Bansal and Shaliastovich 2013, Martin 2013), and in particular with information about long-term risks around the modal Bank of England forecasts driving longer-maturity term premiums in the UK.

These findings point to several important contributions. First, we further the academic debate on the long-run information effect by showing that communication about economic fundamentals alone can move long-term interest rates. We also show that this effect manifests itself particularly via the impact of narrative information on the long-run term premium, which is a novel channel. In other words, we would not necessarily expect a long-run information effect in the absence of narrative information, and the observed movement in long-run rates need not correspond to a change in modal expectations of future central bank policy.

Second, the policy implication of our work is that Delphic forward guidance is alone sufficient to move long-run interest rates. The channel in Hanson and Stein (2015) requires policymakers to generate a change in short-term interest rates in order to move long-term interest rates, but our channel simply requires the transmission of information about economic conditions. We therefore expect it to be effective even in the absence of a change in short-term interest rates or in periods when they are constrained by a lower bound.

Third, we make a methodological contribution by proposing a new test of long-run information in mixed data. The pairing of low-dimensional quantitative information with high-dimensional text data is not unique to the IR. As mentioned above, the Fed releases a textual statement along with its policy rate decision after each FOMC meetings. One could apply our empirical strategy directly to such events to see whether they transmit long-run information.[2] The economic interpretation of such information would be poten-

[2]Gürkaynak et al. (2005) show that the market reaction to Fed statements can be summarized with two factors, one associated with short-run rate movements and another with long-run movements. While they interpret the second factor as a response driven by text, they do not analyze the content of the statement directly. Lucca and Trebbi (2009) show that Fed rate announcements drive short-run rates

tially different to that in the IR, but nevertheless our test of its presence could advance the broader empirical literature beyond this paper.

The remainder of this paper is structured as follows. In Section 2 we provide more detail on the IR and evidence that IR publication days tend to give rise to market news. In Section 3 we present a framework within which to capture why, even at horizons beyond two or three years, communication about the economic situation and outlook might affect interest rate expectations. Section 4 then describes how we convert the IR into quantitative measures to conduct our empirical analysis. Section 5 describes and presents our test of information in the text shocks, while Section 6 refines this by analyzing specific topics in the IR narrative. We then conclude.

2 The Inflation Report and its Impact on Yields

The Bank of England has published a quarterly Inflation Report since 1993, although we focus on the period 1998, after which the Bank had independent control of monetary policy and forecasts were consistently prepared, through mid-2015, after which the Bank of England moved to a different communication protocol for the IR release. During our sample period, the IR was published one week after the contemporaneous policy decision was announced. In total, our sample contains 70 IRs. The IR contains forecast fan charts for GDP growth and inflation summarised by mode, variance and skew. Since 1998 these forecasts have been consistently conditioned on the path for the policy rate (called 'Bank Rate') implied by market interest rates. The IR also contains extensive text data broadly organized into two parts. A set of economics sections assess the current state of the economy, covering recent developments in and the near-term outlook for financial conditions, demand, supply, costs and prices. A forecast section describes the MPC's forecasts, the risks around those forecasts, and the potential trade-offs for policy. The IR does not contain explicit forward guidance, understood as an explicit commitment to a future policy rule. As we will clarify below, however, this does not mean that it does not potentially contain information relevant for long-run rate expectations.

We focus on the IR for several reasons. Since no policy decision occurs on its publication date, and it does not contain formal forward guidance, it allows us to isolate communication solely about the Bank's views on economic conditions. This allows us to interpret the market reaction we observe on IR days directly to *Delphic* communication (Campbell et al. 2012), unlike in other environments where central banks transmit po-

and hawkish sentiment in Fed statements drives longer-run rates. The important part of our test is not the finding that text drives residual variance in asset prices, but that *different* components drive different maturities.

tentially multiple kinds of signals.[3] Second, since the IR is published on a fixed schedule, the timing is not endogenous to evolutions in traders' beliefs nor market conditions more broadly. Third, it contains heterogeneous information in the form of quantitative forecasts and textual narrative, which allows us to explore which type of information drives different market responses.

2.1 Asset Price Data

We consider four nominal rates derived from UK government bond prices: the one-year spot; three-year; five-year; and five-year ahead, five-year forward rates. This allows us to explore the effects of communication at various points on the nominal yield curve. We have collected daily data on these rates from 1 January 1998 through 31 July 2015.[4] The Bank of England's Monetary Policy Committee sets a short-term, nominal interest rate in each month m that we denote i_m. Standard asset-price theory allows us to express our four rates of interest on a given day t as

1. 1-year spot rate:

$$i_{0:12,t} = \frac{i_{m(t)} + \mathbb{E}\left[\, i_{m(t)+1} \mid Z_t \,\right] + \ldots + \mathbb{E}\left[\, i_{m(t)+11} \mid Z_t \,\right]}{12} + \Upsilon_{i_{0:12},t} \tag{1}$$

2. 3-year forward rate:

$$f_{36,t} = \mathbb{E}\left[\, i_{m(t)+36} \mid Z_t \,\right] + \Upsilon_{f_{36},t} \tag{2}$$

3. 5-year forward rates:

$$f_{60,t} = \mathbb{E}\left[\, i_{m(t)+60} \mid Z_t \,\right] + \Upsilon_{f_{60},t} \tag{3}$$

4. 5-year, 5-year rates:

$$i_{60:120,t} = \frac{\mathbb{E}\left[\, i_{m(t)+60} \mid Z_t \,\right] + \ldots + \mathbb{E}\left[\, i_{m(t)+119} \mid Z_t \,\right]}{60} + \Upsilon_{i_{60:120},t} \tag{4}$$

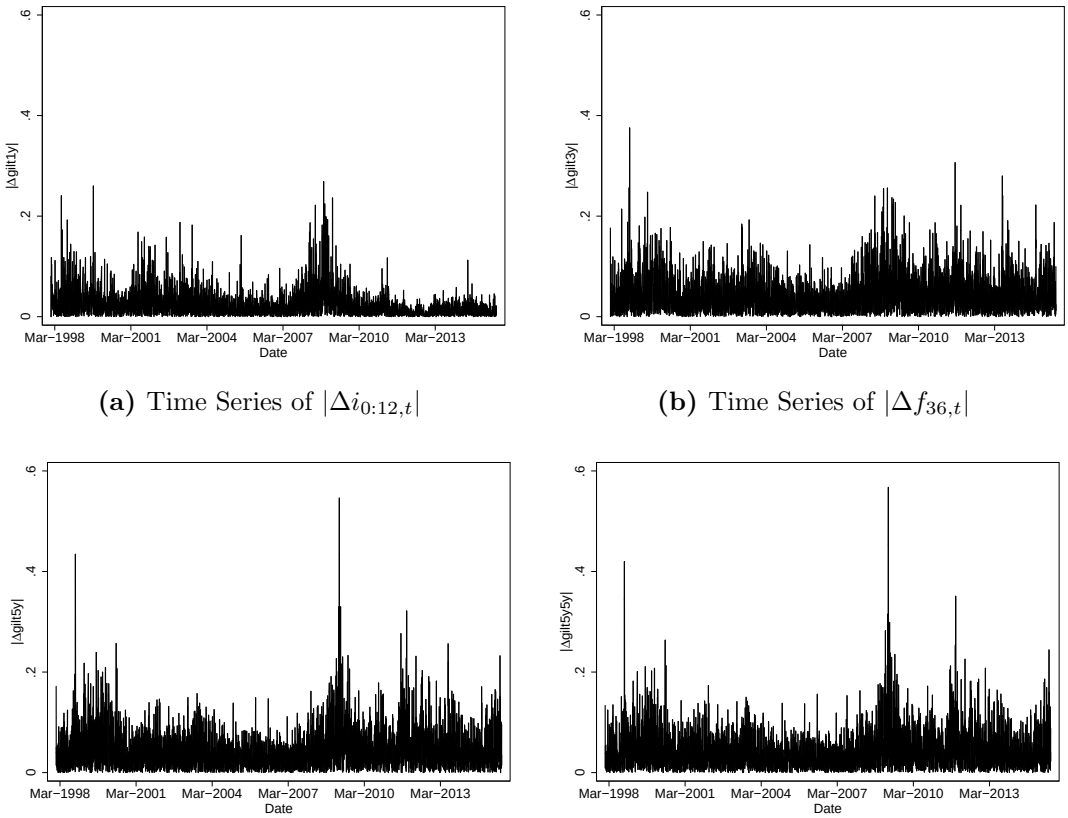
where Z_t is the market information set on day t; $m(t)$ is the month in which day t falls; and $\mathbb{E}$ denotes market expectations.[5] If bond traders were unconcerned about the risks around future interest rates, the term structure of interest rates on UK government

[3]Such an exercise would, for example, not be possible in the US since the Fed does not publish contemporaneous Greenbook forecasts.

[4]The rates we use are those published on the Bank of England's website and are calculated from market bond prices using a variable roughness penalty, spline-based model.

[5]Here we have expressed nominal rates in terms of expectations formed at a monthly frequency for notational convenience; in practice, the forward rates are computed using a notional instantaneous rate of interest, and the 1-year spot and 5-year, 5-year rates are integrals under the curve corresponding to these instantaneous rates.

bonds—the 'yield curve'—should equal the expected path for short-term interest rates. This is often called the 'pure expectations hypothesis' and arises from the ability of traders to choose between buying a long-term bond or investing in a series of short-term bonds. The expected returns from each approach should be equal to rule out arbitrage. In practice, however, market interest rates deviate from the pure expectations hypothesis, with any additional return often referred to as the 'term premium', which we denote by Υ. As explored in Cochrane (2011), there are a number of competing theories for the existence of term premiums and it is likely that they are a driven by a number of factors. In particular, Bansal and Shaliastovich (2013) and Martin (2013) discuss the role of risks and uncertainties around forecasts for growth and inflation in driving term premiums. Given that the IR provides both explicit and implicit forecast distribution information, it is likely that IR information can affect term premiums as well as expectations for short-term interest rates.



(a) Time Series of $|\Delta i_{0:12,t}|$

(b) Time Series of $|\Delta f_{36,t}|$

(c) Time Series of $|\Delta f_{60,t}|$

(d) Time Series of $|\Delta i_{60:120,t}|$

Figure 1: Daily Nominal Rate Moves

Notes: This figure plots the time series of the absolute change in the one-year spot rate; the three-year rate; the five-year rate; and the five-year ahead, five-year rate. These are nominal rates derived from daily Gilt prices from 01/01/1998-31/07/2015.

In our empirical analysis, and in common with much of the literature on central bank communication, we seek to explain absolute daily changes in market rates. Since rates at the beginning of day t already incorporate all information relevant for market expectations, the absolute value of the move on day t is a measure of the news that arrives on the day. Explaining the direction of the rate move requires additional information on changes in implied policy stances. Figure 1 shows the time series for the absolute values of the daily yield movements.

2.2 Event Study

Our primary event of interest is the publication of the IR. While the tendency in the literature is towards using tight, intra-day windows around communication events,[6] we use daily changes since it may take markets longer to fully incorporate the complexity of the IR.

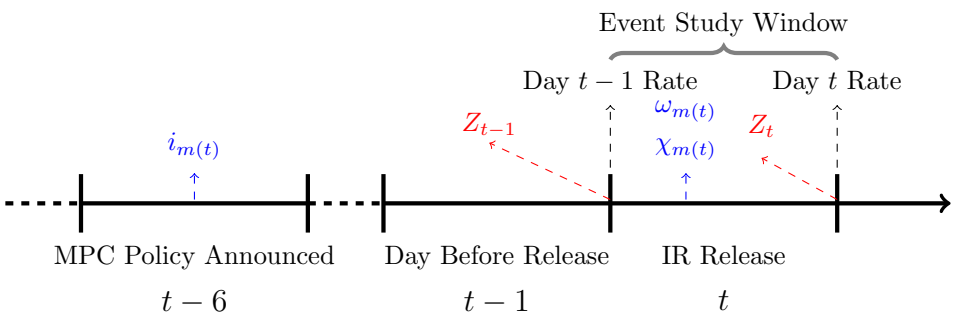


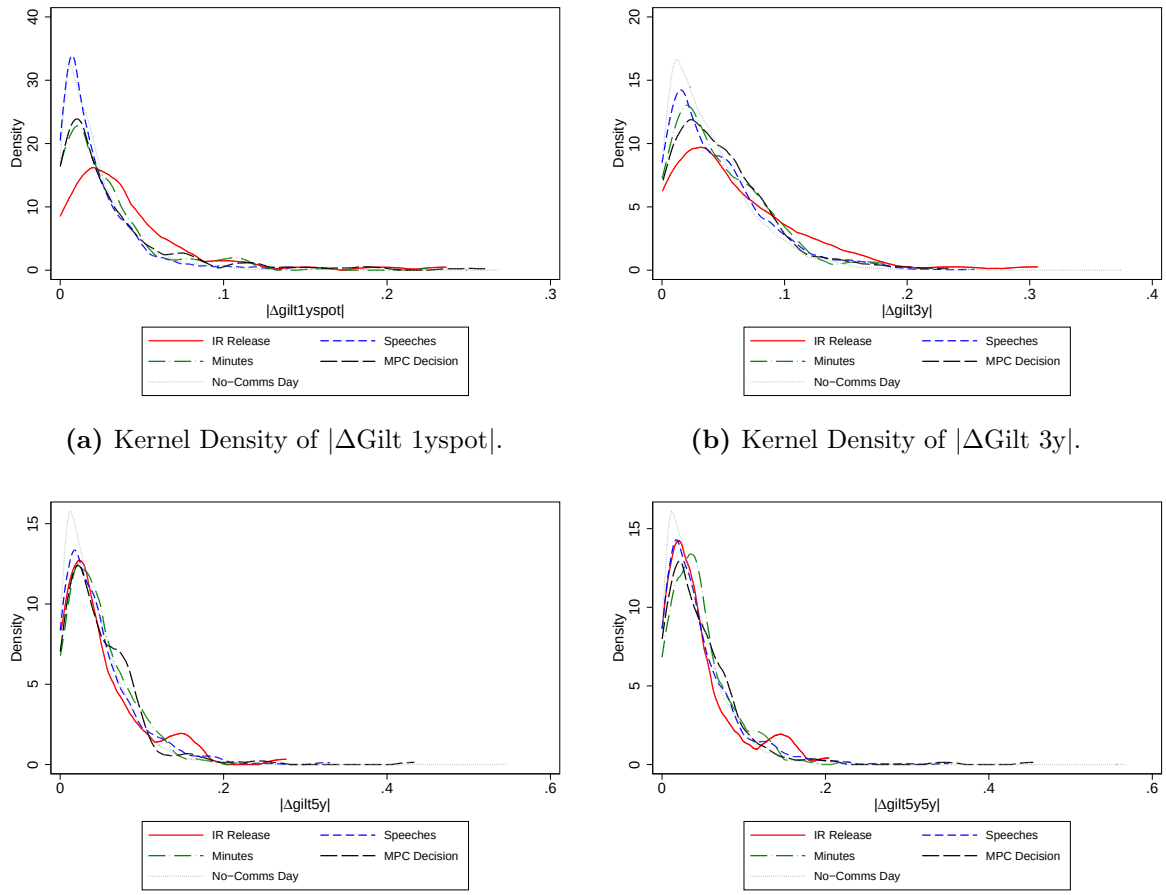
Figure 2: Event Study Time Line for IR Publication on Day t

Figure 2 makes clear the timing of the IR publication on a given day t. Market rates on day $t-1$ incorporate all information available prior to day t, denoted Z_{t-1}. The IR then transmits two kinds of information: quantitative forecast information $\boldsymbol{\omega}_{m(t)}$ and qualitative, narrative information $\boldsymbol{\chi}_{m(t)}$. We denote the IR information by month since its content relates directly to the current month's policy decision $i_{m(t)}$ made one week before day t,[7] but as the next section makes clear, the IR can also impact future rate expectations through several channels. Our identifying assumption is that $Z_t = Z_{t-1} \cup \boldsymbol{\omega}_{m(t)} \cup \boldsymbol{\chi}_{m(t)}$, i.e. all new information relevant for market expectations of future monetary policy on day t is contained in the IR. Given this assumption, we can attribute absolute changes in market rates on IR publication days to the news contained in the IR.

[6] Gürkaynak et al. (2005), Nakamura and Steinsson (2013), and Gertler and Karadi (2015) all use high-frequency identification relying on news about monetary policy in a 30-minute window surrounding scheduled Federal Reserve announcements.

[7] In the week between the announcement of the MPC decision and the IR release, there is not typically monetary news.

Before explaining how IR information can change market expectations, we assess to what extent the IR constitutes market news. We classify each day in our sample according to whether: (1) an IR is released; (2) a policy decision from the MPC is announced; (3) an MPC member makes a public speech; (4) minutes from MPC meetings are released; or (5) none of the above. We then plot kernel densities for each of these five categories in figures 3a-3d. For one-year spot and three-year rates, the IR release dates appear to generate consistently a large amount of news. For longer-dated maturities the effect is less clear cut, but there are large tail moves in interest rates on IR release dates. In appendix A, we conduct a more formal assessment of the market impact of the IR using regression analysis, and continue to find the same patterns as in the kernels.



(a) Kernel Density of |ΔGilt 1yspot|.

(b) Kernel Density of |ΔGilt 3y|.

(c) Kernel Density of |ΔGilt 5y|.

(d) Kernel Density of |ΔGilt 5y5y|.

Figure 3: Kernel Densities of Yield Changes by Type of Communication

Notes: These figures show the kernel-density distribution of changes in expected interest rates at different maturities. We use the Epanechnikov kernel, and set the half-width to the value that would minimize the mean integrated squared error if the underlying distribution were Gaussian.

3 Interpretative Framework for Information Effect

We now present a simple framework within which to think about the potential channels through which the publication of the IR can have an information effect on market rates. First, denote the MPC's month m forecast variables as

$$\boldsymbol{\omega}_m = \left(\mathbb{E}_m^{MPC}[\,\pi_{m+h}\,] - \pi^*, \mathbb{E}_m^{MPC}[\,\tilde{y}_{m+h}\,]\right)^T \tag{5}$$

where $\mathbb{E}_m^{MPC}[\,\pi_{m+h}\,]$ is the MPC's h-month ahead forecast of inflation made in month m, here expressed as a deviation from the inflation target π^*; and $\mathbb{E}_m^{MPC}[\,\tilde{y}_{m+h}\,]$ is the analogous forecast of the output gap. The MPC's policy action in month m can be approximated with a standard monetary policy rule as

$$i_m = r_m^* + \pi^* + \boldsymbol{\phi}^T \boldsymbol{\omega}_m + \epsilon_m. \tag{6}$$

Here r_m^* is the MPC's view of the equilibrium real interest rate, which we allow to be time-varying, albeit slow moving, as in Laubach and Williams (2003); the $\boldsymbol{\phi}$ coefficients relate economic conditions to the nominal rate decision; and the residual ϵ_m is the part of the policy rate that cannot be explained by the central bank's typical reaction to the information set and is, therefore, the common definition of a monetary policy shock as in Romer and Romer (2004) or Cloyne and Hürtgen (2016). We take the $\boldsymbol{\phi}$ coefficients as fixed in the discussion below. In a model with time-varying coefficients, one can think of $\boldsymbol{\phi}$ as the average weights placed on forecasts, and the monetary policy shock as capturing deviations in meeting t from these averages.

one can think of the monetary policy shock as capturing time-variation in these weights

Equations (1)-(4) show that longer-term market interest rates depend on market expectations of future monetary policy. Using (6), we can express the k-month ahead forward rate on day t as

$$f_{k,t} = \mathbb{E}\left[\, r^*_{m(t)+k} \;\middle|\; Z_t \,\right] + \pi^* + \boldsymbol{\phi}^T \mathbb{E}\left[\, \boldsymbol{\omega}_{m(t)+k} \;\middle|\; Z_t \,\right] + \mathbb{E}\left[\, \epsilon_{m(t)+k} \;\middle|\; Z_t \,\right] + \Upsilon_{f_k,t}. \tag{7}$$

This requires the market to predict the MPC's future view of the real interest rate $r^*_{m(t)+36}$; future preferences $\epsilon_{m(t)+k}$; and future *h-period* ahead forecasts $\boldsymbol{\omega}_{m(t)+k}$. Moreover, Martin (2013) shows that, for a fairly relaxed set of assumptions, the term premium will reflect second, third, and higher moments of the distribution around inflation and growth. If we observe a systematic reaction in market interest rates on the day of an IR announcement, then the market must receive information about one of these components. We now describe how quantitative and narrative information can deliver such information.

3.1 Effects of quantitative information

The quantitative information in the IR release on day t is the forecasts $\boldsymbol{\omega}_{m(t)}$ and the distributions around these forecasts. At the time of release, the interest rate decision $i_{m(t)}$ is already known. The most direct impact of the revelation of $\boldsymbol{\omega}_{m(t)}$ on future expectations would be if it were also informative of future forecasts, since (6) makes clear these will affect future interest rates. To make the point more formal, define $\Delta\hat{\omega}_{m(t)+k,t} \equiv \mathbb{E}\left[\,\omega_{m(t)+k} \mid Z_t\,\right] - \mathbb{E}\left[\,\omega_{m(t)+k} \mid Z_{t-1}\,\right]$. A direct effect on the k-month ahead forward rate is present if $\Delta\hat{\omega}_{m(t),t}$ (the market's updated belief on the current forecast variables induced by the IR publication) is correlated with $\Delta\hat{\omega}_{m(t)+k,t}$. While persistence in macroeconomic conditions might generate this correlation at shorter horizons, it is less likely to generate significant correlations beyond one ($k = 12$) or two ($k = 24$) years, a point we return to in the next section. A similar point can be made with respect to the distributional information around the forecasts. While this in principle can affect future term premiums, whether it is persistence enough to do so is unclear.

Quantitative information can also have important indirect effects. The policy rule (6) presents the market with an identification problem. While the output $i_{m(t)}$ is known prior to day t, the inputs that led to the policy rate are unknown. A higher-than-expected rate might be due to the MPC's having higher beliefs on equilibrium real rates, higher inflation or growth forecasts, or having a more hawkish tilt in its policy stance. The publication of the forecasts eases this problem. Holding fixed a value for $r^*_{m(t)}$, knowledge of $\boldsymbol{\omega}_{m(t)}$ allows the market to back out $\epsilon_{m(t)}$ and vice versa. Furthermore, both equilibrium real rates and policy stances can be highly persistent. The real interest rate is often modeled as a unit-root process, which implies that any revisions to current views on its value would propagate one-for-one into future views. Also, for example, if a more-hawkish-than-expected tilt coincides with the appointment of new members to the MPC, the market might infer a more hawkish stance would persistent throughout their terms, which are three or more years.[8]

3.2 Effects of narrative information

The narrative information in the IR can also have important effects on future expectations for several reasons. First, there are many hundreds of hard and soft indicators of economic activity that the MPC regularly monitors, including surveys, disaggregate activity and inflation series, and information from regional agents. These indicators are all (potentially) endogenously related to each other and to the inflation and output fore-

[8] In our sample period, minutes containing individual MPC members' votes were published around two weeks after the IR, so any inference made on the policy stance due to forecast variable publication must be made for the MPC as a whole.

casts contained in $\boldsymbol{\omega}_{m(t)}$. The narrative in the IR provides the Bank of England's views about the nature of these endogenous relationships, as well as what are the key drivers of the current forecasts. This can influence market views of likely future MPC forecasts. For example, the IR can reveal whether the inflation forecast is driven by persistent or transitory price movements.

Second, along similar lines, the narrative can provide context about the variance and skew around the quantitative forecasts that can affect term premiums in forward rates. Moreover, as we discuss below, in practice there is very little change in the variance of forecasts in our sample, so these may not be accurate signals of the uncertainty facing the MPC about future economic conditions. The narrative can therefore guide markets about the risks the MPC considers when making its forecasts.

Third, neither the MPC nor monetary policymakers more generally typically publish quantitative views on the value of latent macroeconomic variables such as the equilibrium real interest rate. While an important driver of the policy action, r_m^* is an inherently elusive variable that depends on quantities such as the unobserved productive capacity of the economy about which there may be significant disagreement. In this context, the narrative may be the *only* way the MPC can signal its view. As discussed above, if markets understand that the MPC has revised its view on real rates, this can have long-run effects on forward rates. The narrative can also communicate uncertainty about the level of the real rate, which would propagate into long-run term premiums.

A final point is that the information content of the quantitative and narrative signals is not in general separable. For example, the narrative can provide a signal on r_m^*, which, when combined with the published $\boldsymbol{\omega}_{m(t)}$, can fully overcome the identification problem that (6) presents markets. In our empirical strategy, we purge asset price movements *and* narrative information of their variation due in the quantitative forecasts in order to isolate the impact of narrative information.

4 Measuring Inflation Report Content

We now describe how we construct measures of the quantitative and narrative components of the Inflation Report.

4.1 Quantitative information

Table 1 shows the coefficients from estimating a simple AR(1) process on the forecast variables released in the IR over our sample period. Only the variance terms are persistent and actually these variables are rarely changed making them easiest to forecast.

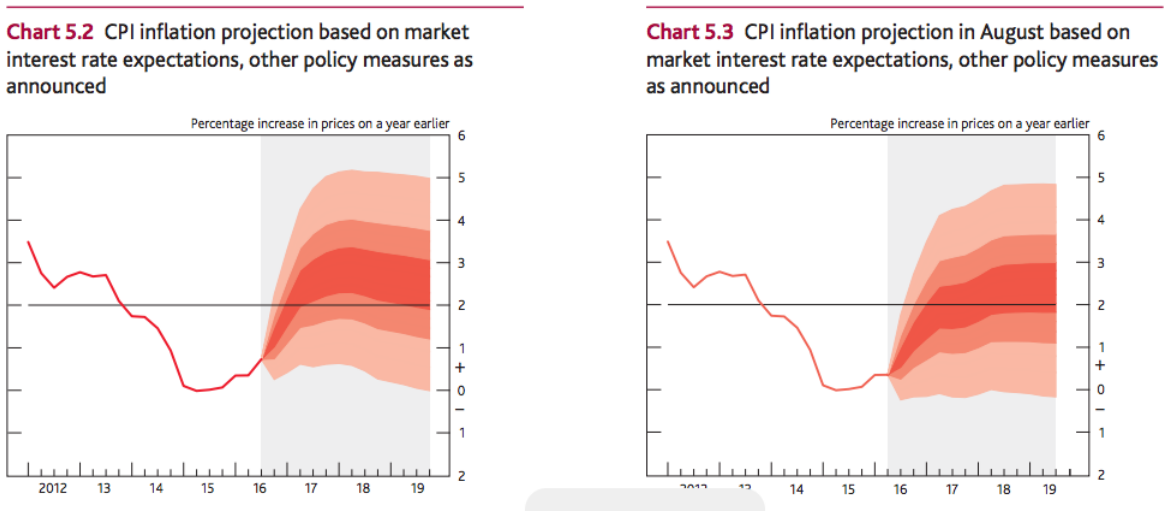


Figure 4: Quantitative Information: Fan Charts

Notes: The left-hand figure shows the fan chart associated with the CPI inflation forecast in the November 2016 Inflation Report. The right-hand figure shows the chart for the August 2016 Report.

In each IR, the Bank published fan charts for the evolution of GDP growth and inflation over the following two years and, from August 2004 onwards for the following three years. Figure 4 provides an illustration of such a fan chart. The Bank also provides quantitative values of the mode, variance and skew of the distributions shown in the fancharts. Those three moments pin down the distribution. We use these values for the two-year horizon as that is the horizon that has tended to be focused on in the Bank's monetary policy communication as the one most relevant for the current stance of policy. We also use a measure of the difference from market expectations for modal variables, and the change from the previous IR release for the variance and skew variables. In order to map more closely with the monetary policy rule in (6), we also include the implied output gap and its change. Altogether we then have fifteen numerical variables from the IR. We add the VIX index as an additional control as a proxy for general market volatility that may be affecting interest rates on the day of the IR release.

4.2 Narrative information

The narrative information in the IR is contained in text. In the 70 Reports in our sample, there are 15,023 distinct paragraphs. We first pre-process the data by removing all non-alphabetic terms, as well as extremely common words that are uninformative about content such as 'the', 'and', and so on—so-called *stopwords*. We then stem each remaining term into its linguistic root using the Porter stemmer. Stems need not be English word: the stem of 'inflation' is 'inflat'. Following these steps, we have 754,884

Main Regressors	(1) π^f_{t+2}	(2) $\hat{y}^f_{t+2}$	(3) $\text{Var}(\pi^f_{t+2})$	(4) $\text{Var}(\hat{y}^f_{t+2})$	(5) $\text{Skew}(\pi^f_{t+2})$	(6) $\text{Skew}(\hat{y}^f_{t+2})$
AR1 Coefficient	0.615***	0.590***	0.869***	0.788***	0.502***	0.628***
D(Crisis)	-0.118*	0.0685	0.240**	0.371***	0.0304	-0.0954
Constant	-0.00490	1.075***	0.0611*	0.255*	0.00982	-0.0317
R-squared	0.532	0.370	0.977	0.870	0.291	0.514
After 1 year	0.14	0.12	0.57	0.38	0.06	0.16
After 3 years	0.00	0.00	0.18	0.06	0.00	0.00
After 5 years	0.00	0.00	0.06	0.01	0.00	0.00

Table 1: Persistence of Key Forecast Variables

total terms in the data and 4,382 unique terms.

Our first representation of the IR text data uses not individual terms, but adjacent two-term phrases—also called *bigrams*—such as 'slow growth' and 'strong growth'. There are 192,753 unique bigrams in the pre-processed data. We drop any bigram that appears in five or fewer Reports, which leaves 22,211 unique bigrams. We then count the frequency of each bigram across each Report. Conceptually, this yields a $70 \times 22,211$ *document-term matrix* whose (t, v)th element is the count of bigram v in the Inflation Report released on day t. This is a high-dimensional object in the sense that the number of dimensions of variation (i.e. the number of bigrams) across Reports exceeds by an order of magnitude the number of Reports in our sample.

The second representation uses a probabilistic topic model called Latent Dirichlet Allocation (LDA), first used in the economics literature by Hansen et al. (2017). Here we provide a high-level overview of LDA; estimation follows the same Markov Chain Monte Carlo procedure described in Hansen et al. (2017) and introduced by Griffiths and Steyvers (2004); we refer interested readers to those papers for full details.

LDA is a Bayesian factor model for discrete data. Suppose there are D documents (we treat each paragraph as a document, so $D = 15,023$) that comprise a corpus of texts with V unique terms (for LDA we return to using individual terms, so $V = 4,382$). The first important objects in LDA are K *topics* (i.e. factors), each of which is a probability vector $\boldsymbol{\beta}_k \in \Delta^{V-1}$ over the V unique terms in the data. The choice of probability distributions is important since it allows the same term to appear in different topics with potentially different weights. Informally, one can think of a topic as a weighted word list that groups together words that all express the same underlying theme.

LDA is a mixed-membership model in which each document can belong to multiple topics. Formally, this is represented by each document d having its own distribution over

topics given by $\boldsymbol{\theta}_d$ (i.e. factor loadings). Informally, θ_d^k represents the "share" of topic k in document d.

The probability that any given word in document d is equal to the vth term is therefore $p_{dv} \equiv \sum_k \beta_k^v \theta_d^k$ and the overall likelihood is $\prod_d \prod_v p_{d,v}^{n_{d,v}}$ where $n_{d,v}$ is the number of times terms v appears in document d. Importantly, LDA reduces the dimensionality of each document substantially. In the document-term matrix, documents live in a V-dimensional space. After estimating LDA, one obtains a representation of each document in terms of the (estimated) θ_d, which lives in the $K-1$ simplex, and in general $K \ll V$. Importantly, though, LDA does not ignore any dimensions of variation in the raw term counts since the underlying topics are free to lie anywhere in the $V-1$ simplex.

LDA places Dirichlet priors over the $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ probability vectors, and the inference problem is to approximate their posterior distributions. The main model selection choice is the number of topics K. We use a model with $K = 30$, which provides a generally interpretable latent space.[9] Figure 5 represents the 30 topics that LDA estimates in our data. An initial observation is that topics are indeed interpretable: topic 6 appears to capture discussion of commodity prices; topic 14 of the forecast; topic 24 of financial markets; and so on. Since topics have no natural ordering, we define our own based on whether an IR is published during a cycle of rate increases (i.e. the previous rate change was an increase) or rate decreases (i.e. the previous rate change was a decrease). For each topic, we compute its average share of time in the IR during both cycle, and order topics based on the difference: topic 0 (29) is most associated with an increasing (decreasing) cycle. Again, this gives interpretable results. For example, topic 0 captures concerns about wage growth, and topic 29 reflect financial market conditions, which were of primary concern during the crisis.

While we estimate LDA at the paragraph level to exploit variation across thousands of examples of text, we are ultimately interested in the content of entire Reports. We follow the procedure detailed in Hansen et al. (2017) to obtain these given the estimated LDA model. We denote by $\boldsymbol{\theta}_t$ the distribution over topics in the IR published on day t. Since changes in topic coverage can also have potentially important market effects, we define the vector $\boldsymbol{\delta}_t \equiv \boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}$ and thereby obtain a sixty-dimensional representation of each text—the 30 topic levels and 30 topic changes.

Figure 6 plots the share of time the Reports in our sample spend covering two illustrative topics for which we provide an alternative representation in terms of word clouds. Topic 15 reflects discussion of labor markets. This was fairly stable until 2014, when

[9] There is a well-known trade-off between interpretability and goodness-of-fit in the machine learning literature (Chang et al. 2009). While objective measures of goodness-of-fit can be used to determine a choice for K, our goal is to obtain an interpretable description of IR content, for which defining objective criteria is challenging.

													Pro-cyclicality
Topic0	growth	wage	cost	earn	pay	pressur	labour	sector	real	privat	higher	rise	0.013
Topic1	measur	chang	use	differ	can	weight	indic	rel	provid	movement	base	tend	0.013
Topic2	rate	interest	market	offici	expect	impli	point	chart	rise	path	unchang	forward	0.013
Topic3	growth	quarter	gdp	slow	output	weaker	slowdown	strong	stronger	trend	grew	pick	0.01
Topic4	survey	sector	output	servic	suggest	manufactur	indic	report	evid	busi	agent	region	0.01
Topic5	invest	capit	busi	money	financi	deposit	compani	sector	corpor	broad	increas	privat	0.009
Topic6	price	oil	energi	rise	increas	cost	commod	higher	import	food	futur	ga	0.009
Topic7	month	year	rose	annual	fell	point	percentag	earlier	chart	around	increas	januari	0.007
Topic8	import	uk	export	trade	good	net	demand	domest	servic	sterl	account	world	0.006
Topic9	report	sterl	februari	exchang	august	rate	novemb	start	lower	time	point	averag	0.006
Topic10	estim	data	revis	gdp	suggest	account	recent	output	previous	past	inform	current	0.005
Topic11	household	spend	hous	consumpt	incom	consum	real	increas	save	rise	wealth	ratio	0.005
Topic12	may	recent	retail	level	year	sale	stock	reflect	time	underli	trend	inventori	0.004
Topic13	remain	year	past	recent	level	although	low	continu	littl	broadli	appear	risen	0.003
Topic14	project	period	forecast	central	mpc	assum	assumpt	path	around	next	report	profil	0.002
Topic15	employ	labour	unemploy	work	number	hour	market	job	peopl	increas	time	rise	0.0
Topic16	may	could	reduc	might	exampl	increas	possibl	adjust	lead	greater	addit	becom	-0.001
Topic17	chart	year	averag	tabl	probabl	show	around	forecast	distribut	rang	outturn	time	-0.001
Topic18	risk	committe	uncertainti	outlook	balanc	judg	view	central	downsid	prospect	overal	consider	-0.004
Topic19	effect	will	govern	impact	increas	public	plan	fiscal	year	nomin	chang	budget	-0.004
Topic20	inflat	target	cpi	pressur	price	rpix	higher	vat	inflationari	rise	contribut	downward	-0.004
Topic21	expect	term	near	littl	month	medium	short	outlook	rise	next	longer	fall	-0.006
Topic22	will	section	factor	depend	discuss	develop	extent	suppli	box	affect	impact	demand	-0.007
Topic23	economi	unit	euro	area	global	world	kingdom	activ	state	countri	prospect	econom	-0.008
Topic24	market	equiti	financi	asset	bond	yield	corpor	share	may	increas	return	reflect	-0.008
Topic25	fall	reflect	declin	chart	part	fallen	sharpli	larg	past	lower	fell	much	-0.008
Topic26	demand	continu	growth	recoveri	remain	domest	weak	activ	support	pace	gradual	eas	-0.011
Topic27	bank	polici	committe	purchas	mpc	meet	asset	monetari	rate	maintain	billion	econom	-0.015
Topic28	product	compani	capac	margin	suppli	recess	spare	demand	weak	economi	output	slack	-0.018
Topic29	bank	credit	condit	lend	fund	borrow	mortgag	loan	avail	compani	secur	cost	-0.02

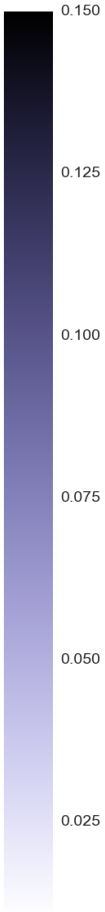


Figure 5: Topics Ranked by Pro-Cyclicality; Terms within Topics Ranked by Probability

Notes: This table summarizes the 30 separate distributions over vocabulary terms that LDA estimates to represent topics. We order these distributions from 0 to 29 based on a pro-cyclicality index that computes the difference in average time the IR spends discussing the corresponding topic when interest rates are in tightening and loosening cycles, respectively. Within each row, terms are ordered left to right by the probability they appear in each topic, with differential shading indicating approximate probability values.

the Bank started increasing its analysis of the labor market in response to the puzzle that domestic inflationary pressure had remained subdued even as unemployment fell. Topic 26 about demand had a marked increase at the onset of the financial crisis, and has remained high along with the Bank's concerns about the pace of its recovery.

(a) Topic 15 'Labor Market' **(b)** Topic 26 'Demand'

Figure 6: Illustrative Topic Variation across Inflation Reports

Notes: These figures plot the prevalence of two illustrative topics in the Inflation Reports in our sample. Recession periods are shaded in gray. The distributions over terms that each topic induces are represented as word clouds, where the size of term is approximately proprtional to its probability.

These distinct representations of text are useful for different reasons. The bigram counts provide a detailed description of each Report, and come closer to capturing the full amount of information available to markets. On the other hand, the counts' sheer dimensionality makes them an unwieldy means of summarizing the narrative and for comparing content across reports, which are strengths of LDA. For this reason, we use both in our empirical analysis.

5 How significant is the information effect?

5.1 Impact of the quantitative information

The first step in our analysis is to take the numerical information from the IR forecast, together with the VIX uncertainty index, and use it to purge the effects of numerical data on the asset price news. This regression is done in 2. Given the correlation between some of the forecast variables, and in particular the correlations of the changes in many of the variables, we are not (particularly) interested in the significance of any individual coefficients. Rather, we wish to purge the effect of the whole selection of the numerical data. As such we focus on the joint explanatory power and look at the R-Squared for each of the three maturities.

The key take away is that, consistent with expectations theory of the yield curve and a standard Taylor-type monetary policy rule, the two-year ahead forecasts play a greater role in explaining the news in shorter-maturity interest rates more than longer maturity yields. The two-year ahead forecast numbers still provide some information about economic conditions at longer maturities, but effect declines with time. We denote by $\hat{\nu}_t^H$ the residual of the forecast variable regression.

5.2 Does the narrative contain additional information?

Our main empirical challenge is to determine whether there is information contained within the text shocks is relevant for explaining the asset price news residual $\hat{\nu}_t^H$. In other words, we ask whether the variation in text that is not driven by the forecast variables can explain the variation in market rates that cannot be explained by them.

The problem is that one clearly cannot simply regress $\hat{\nu}_t^H$ on all text shocks since there are 12,000 regressors and less than 70 degrees of left to explain given $\hat{\nu}_t^H$ is itself a residual from a regression with 17 controls and 70 observations. Moreover, one cannot simply correlate individual text shocks with $\hat{\nu}_t^H$ since even in a randomly generated dataset there will be some significant but spurious relationships when one has available 12,000 covariates.

To solve this issue, we first fit a LASSO model (Tibshirani 1996) of the form

$$\sum_t \left(\hat{\nu}_t^H - \boldsymbol{\beta}' \cdot \boldsymbol{\zeta}_t\right)^2 + \lambda \sum_v |\beta_v| \tag{8}$$

The LASSO adds an L1 penalty term in the regression coefficients to the residual sum of squares. As one increases the weight on the penalty, given by λ, more and more of the estimated regression coefficients become zero. Thus LASSO has become a popular tool

Main Regressors	(1) $\lvert\Delta 1y\rvert$	(2) $\lvert\Delta 3y\rvert$	(3) $\lvert\Delta 5y\rvert$	(4) $\lvert\Delta 5y5y\rvert$
π^f_{t+2}	-0.016	-0.014	0.016	0.025
	[0.535]	[0.698]	[0.645]	[0.432]
$\lvert\Delta\pi_t^{MPC,+8} - \Delta\pi_t^{P,+8}\rvert$	0.034	-0.035	-0.042	-0.051
	[0.292]	[0.442]	[0.314]	[0.152]
$\mathrm{Var}(\pi^f_{t+2})$	-0.011	0.025	0.021	0.020
	[0.324]	[0.161]	[0.257]	[0.218]
$\Delta\mathrm{Var}(\pi^f_{t+2})$	0.038	-0.057	-0.12**	-0.085*
	[0.455]	[0.240]	[0.011]	[0.084]
$\mathrm{Skew}(\pi^f_{t+2})$	0.0045	0.0100	-0.0040	-0.0036
	[0.891]	[0.834]	[0.924]	[0.923]
$\Delta\mathrm{Skew}(\pi^f_{t+2})$	0.025	0.065	0.068	0.052
	[0.500]	[0.150]	[0.108]	[0.154]
$\hat{y}^f_{t+2}$	0.031**	0.00010	-0.012	-0.015
	[0.038]	[0.995]	[0.476]	[0.381]
$\lvert\Delta y_t^{MPC,+8} - \Delta y_t^{P,+8}\rvert$	-0.048*	0.050	0.053*	0.024
	[0.063]	[0.140]	[0.084]	[0.373]
$\mathrm{Var}(\hat{y}^f_{t+2})$	0.0010	-0.0052	-0.0021	-0.012
	[0.934]	[0.765]	[0.903]	[0.446]
$\Delta\mathrm{Var}(\hat{y}^f_{t+2})$	0.0098	0.019	0.024	0.034*
	[0.539]	[0.236]	[0.153]	[0.058]
$\mathrm{Skew}(\hat{y}^f_{t+2})$	-0.014	-0.071	-0.060	-0.045
	[0.688]	[0.171]	[0.202]	[0.260]
$\Delta\mathrm{Skew}(\hat{y}^f_{t+2})$	-0.063**	-0.00055	0.0031	-0.022
	[0.049]	[0.990]	[0.938]	[0.533]
$y,\tilde{+}8_t$	0.41	3.35*	1.68	-0.50
	[0.722]	[0.097]	[0.369]	[0.747]
$\lvert y,\tilde{+}8_t - y,\tilde{+}8_{t-1}\rvert$	4.96**	1.41	-1.00	-1.43
	[0.015]	[0.685]	[0.767]	[0.595]
$\lvert y,\tilde{+}7_t - y,\tilde{+}8_{t-1}\rvert$	1.64	-0.35	-0.25	1.31
	[0.430]	[0.925]	[0.944]	[0.652]
VIX	0.0013**	0.0023	0.0018	0.0017
	[0.023]	[0.151]	[0.229]	[0.150]
Constant	-0.070*	-0.026	0.014	0.043
	[0.075]	[0.693]	[0.833]	[0.456]
Observations	69	69	69	69
R-squared	0.563	0.368	0.280	0.274
R-squared No Vix	0.526	0.303	0.229	0.215

Table 2: Purging $\lvert\Delta i_t^H\rvert$ of forecast variables

for model selection in economics (see Belloni et al. 2014).

A major challenge with LASSO is the appropriate selection of λ, and several of its theoretical asymptotic properties depend on this. In finite samples, a popular default is to choose λ via cross validation, which targets out-of-sample predictive performance. Given our small sample size, we opt for leave-one-out cross validation (LOOCV). The specific algorithm is:

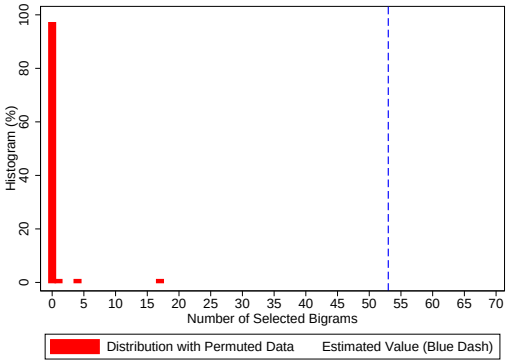
1. For each of a sequence of possible λ penalty coefficients:
   (a) For each of the 69 data points:
      i. Remove the point from the sample.
      ii. Fit (8) on the remaining 68 points.
      iii. Calculate the forecasted value for the held-out point from the fitted model, and compute the squared error.
2. Select λ to the highest that has a MSE still within 1 s.d. of the MSE-minimising λ across 69 out-of-sample forecasts.

The MSE-minimising λ and the number of retained bigrams are presented in table **??**.

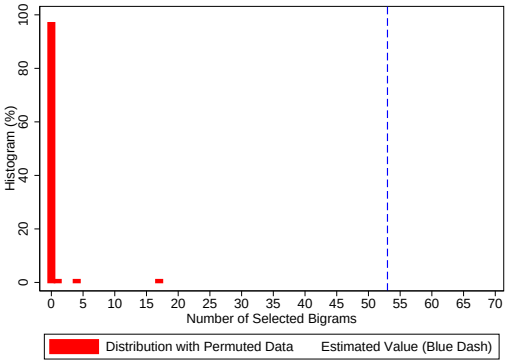
	Gilt1yspot	Gilt3y	Gilt5y	Gilt5y5y
λ selected	0.000428	0.000899	0.000837	0.000668
Bigrams selected	60	57	53	55

Table 3: Number of Selected Text Features from LOOCV

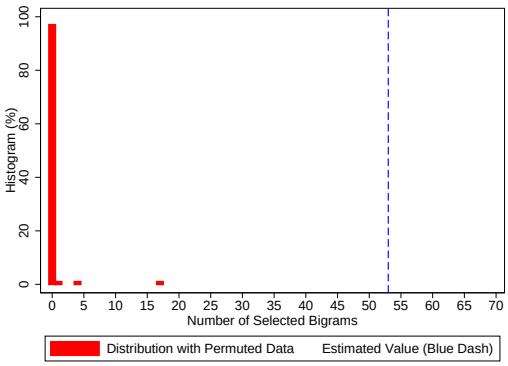
The fact that many topic features are retained in the LOOCV exercise is not *primae facie* evidence that topics contain news because we do not have a well-defined null hypothesis to test against. To do so, we simulate the distribution of the number of selected features when bigrams have a random relationship to the interest rate residuals. In each of 100 simulations, we randomly permute the dependent variable in the LASSO, repeat the LOOCV procedure, and store the number of selected bigrams. We then compare the values it table 3 against these simluated distributions. The results are contained in figure 7, where the red-shaded histogram is the distribution of the number of selected bigrams under the null of random relationships, and the blue dashed line is the number of bigrams we select in our actual data. At all maturities we strongly reject the null hypothesis that the correlations we find are spurious, and we conclude that indeed there is genuine explanatory power contained in the IR that is orthogonal to forecast variables.



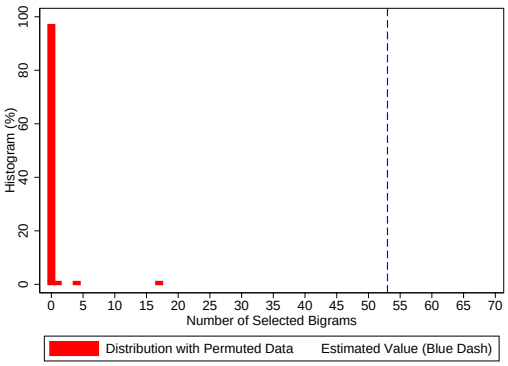
(a) 1-year spot rate



(b) 3-year Forward



(c) 5-year Forward



(d) 5-year, 5-year rate

Figure 7: Novel Test for News in Textual Data

6 What are the channels of the information effect?

6.1 Long-run versus short-run effects

Our test above indicates that, even after purging the effect of the quantitiative signal ω_t, the IR contains news relative for both short and longer-term market interest rates. The question then becomes: through what channels is the IR information affecting the yield curve. One way to shed light on this is to examine what text content is most relevant for explaining the moves in yields and whether that changes depending on the horizon of the yield in question. If it is the same content affecting both short and long-term yields then this could suggest that the long-run effect is purely a mechanical consequence of the short-term move rather than there being any information content that is actually relevant for the long-run. In contrast, if different text content is driving the long-run effect then this could suggest there are some types of information that are more relevant for affecting long-term expectations or perceptions of long-term risks.

While we could examine the bigrams selected as part of the LASSO procedure above, the maximum number of relevant covariates LASSO can select is limited by the sample size. It is likely, therefore, that a larger set of IR information is important for market interest rates than just those bigrams selected above by the LASSO procedure.

To explore this question, we shift to examining the information content of the estimated topic distributions of the IR. In other words, whether there is information contained within $\boldsymbol{\theta}_t^R$ and $\boldsymbol{\delta}_t^R$ that is not driven by the numerical forecast variables and is relevant for the residual variation in market rates $\hat{\nu}_t^H$.

The results for the information test carried out on the topic representations of the text are reported in the appendix. Suffice it to say that the use of LDA does not affect the conclusion that the IR represents an information shock for markets. However, even with the reduced dimensionality we still struggle to identify the core topics. This is because, as is well known in the statistics literature, the number of selected features from a LASSO estimated via cross validation may be a superset of the relevant variables (Meinshausen and Bühlmann 2006). That is, LASSO screens relevant variables so that their inclusion is guaranteed in the selected set (at least asympototically) but does not necessarily select the true model in the sense that noise variables are also guaranteed to be in the selected set.

In general the problem of post-selection inference for LASSO estimates is quite challenging, and not fully resolved in the statistics literature. The conditions required for model selection consistency—i.e. the conditions under which one can recover exactly the set of relevant variables from the LASSO—are quite stringent in theory and difficult to verify in finite samples.

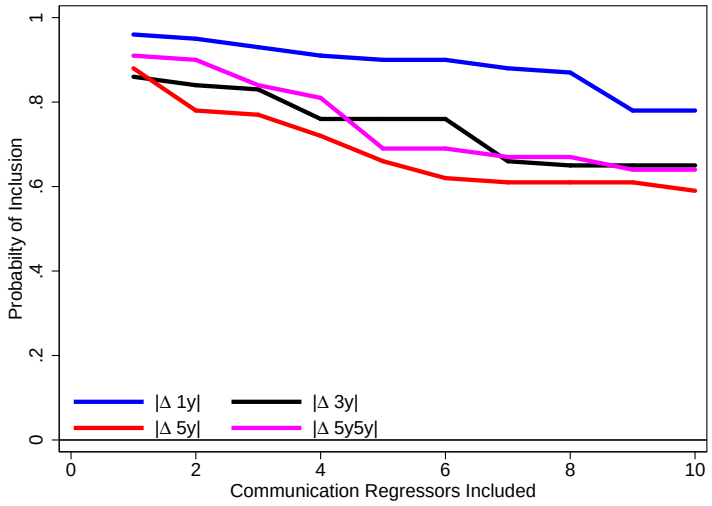
As an alternative to formal statistical inference, we instead follow a bootstrap procedure suggested by Hastie et al. (2015). In each of 500 simulations, we draw a bootstrap sample with replacement from our original data, compute coefficient estimates using the LOOCV procedure detailed above, and record whether each variable is selected. Thus across all the bootstrap draws, we can compute the fraction of times that each topic variable selected, and use this as a guide to which are the key variables driving the market response to the IR. This is closely related to the stability selection approach of Meinshausen and Bühlmann (2010), who provide formal bounds on the probability of false inclusion as a function of the fraction of times a variable appears in random subsamples.

After estimating the bootstrap draws, for each maturity of interest, we regress $\hat{\nu}_t^H$ on, respectively, the M topics with the highest estimated inclusion probabilities. Figure 8a plots the marginal probability of inclusion of the next highest text variable as we add more and more text regressors. One thing that stands out is that the probability of inclusion declines less quickly for the short end of the yield curve. This suggests that a broader range of the topics contain some useful information for the short end.

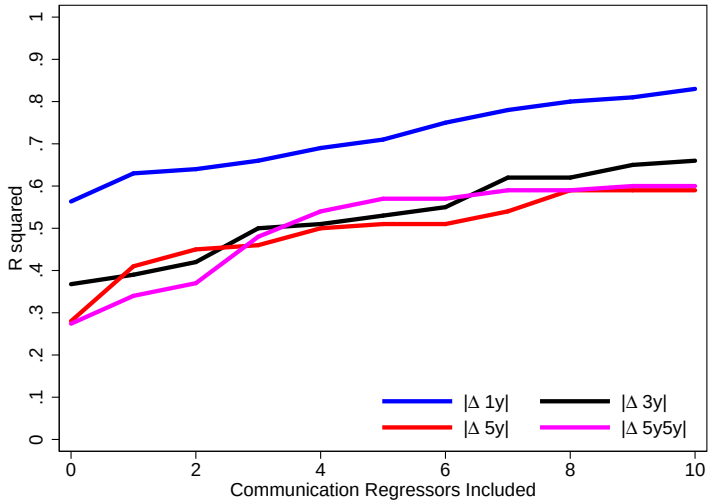
In each of these regressions, we can then compute the R^2 from these regressions and thereby obtain a measure of the relative amount of variation in market rates that we can attribute to χ_t as we include more of the text data. The plots of the rolling-R^2 statistics are in figure 8b. As already discussed, the amount of variation in market yields that the quantitative signals explain is greatest for the one-year spot rate. The relative contribution of text variables, therefore, is weakest for one-year spot rates, stronger for three-year rates, and strongest for five-year and five-year ahead, five-year rates (these last two rates are essentially indistinguishable). For the longest maturities, including around five topic controls is already enough to capture as much variation in market rates as all of the forecast variables.

In figure 9 and 10 we present the top 4 topics included in the regressions on the one-year spot rate and the five-year, five-year forward rates respectively. The most striking aspect of this is that the topics are clearly different. This indicates that it is different dimensions of the inflation report text that explains the market news. In particular, the topics explaining the short-end reaction cover more conjunctural economic news while the topics explaining the long-end reaction are more to do with forecasts, risks around it and the Bank's inflation forecast.

To formalise this finding, Table 4 shows the Spearman rank correlations of topic features across asset classes. That is, we rank each of the topic features by their probability of inclusion for each asset class and then compare the correlation of ranks across assets. The topics driving the longer-maturity asset classes are very similarly ranked. The striking result is that the topics that drive the one-year spot yield are ranked very different



(a) Effect of adding more controls: Prob(Inclusion) - Nominal Yields

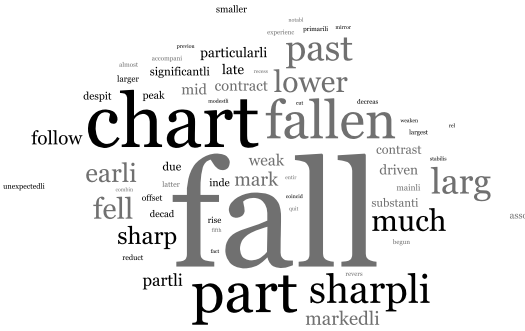


(b) Effect of adding more controls: R^2 - Nominal Yields

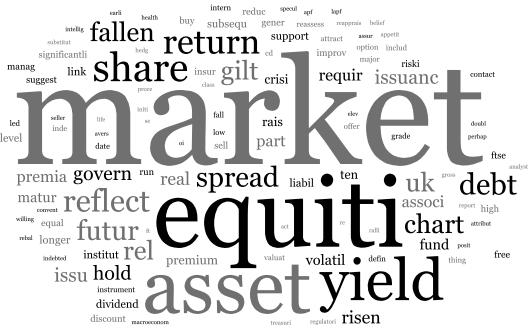
Figure 8: Prob(Inclusion) and R^2 expained by the text topics

Notes: XXXX.

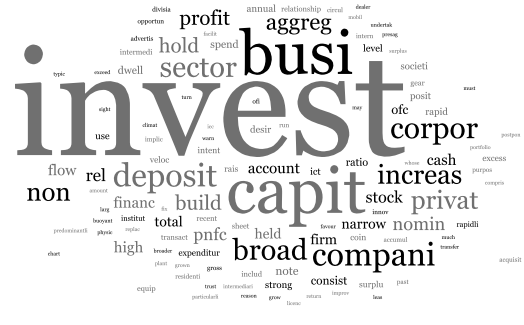
Figure 9: Post-Selection Topic Evaluation: 1 year spot



(a) T29



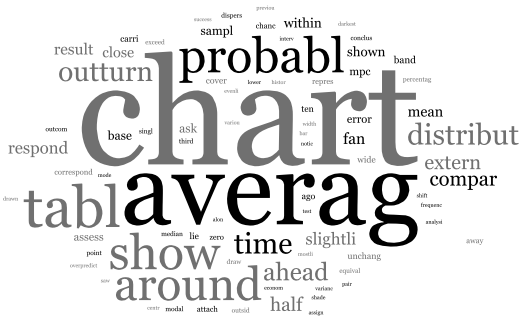
(b) D7



(c) T18

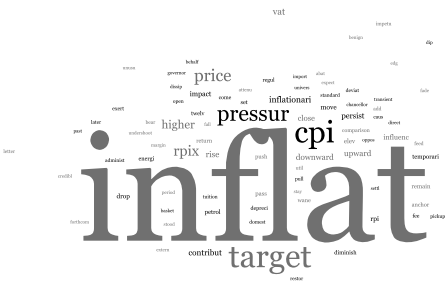
(d) T8

Figure 10: Post-Selection Topic Evaluation: 5y5y

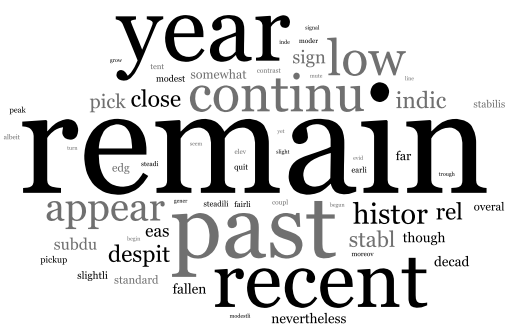


(a) D4

(b) D20



(c) T6



(d) D12

from other yields. Table 5 shows the top and bottom topics, with their probability of inclusion, for the shortest and longest maturity assets. The ranking is almost perfectly orthogonal.

Table 4: Spearman rank correlations of topic features

	1yspot	3y	5y	5y5y
1yspot	1	·	·	·
3y	0.19	1	·	·
5y	0.05	0.68***	1	·
5y5y	0.11	0.47***	0.82***	1

Table 5: Top and bottom topic features for 1-year Spot and 5y5y Rates

Gilt1yspot		Gilt5y5y	
T29	0.958	D4	0.91
D7	0.954	D20	0.896
T18	0.935	T6	0.836
T8	0.91	D12	0.808
...	...	...	...
D20	0.528	T18	0.636
D4	0.47	D7	0.558
D12	0.32	T8	0.322
T6	0.192	T29	0.322

Taken altogether, we have found compelling evidence that text provides information on long-term economic conditions not captured by the forecast variables and that this drives as least part of the move in long-term market interest rates.

6.2 Expectations and Term Premiums

Another way to shed light on the channels behind the information effect observed is to examine the effects on expectations and term premiums seperately. As described in section **??**, information influencing market expectations of future real interest rates or inflation would be expected to affect the expectations component, whereas information on the risks around the outlook would be expected to affect the term premium.

One common way to decompose changes in yields into expectations and term premiums is to use an affine term structure model. Some of these models use only the past behaviour of the market yield curve to estimate this decomposition, whereas others supplement that with survey or other additional data on expectations. The specification of

Table 6: Post-Selection Topic Evaluation: 1 year spot

	Total Var	Var(Exp)	Var(TP)	2 x Cov
1 year spot	0.0032	0.0024	0.0001	0.0007
	100	75	3	22
3 year	0.0066	0.0037	0.0009	0.0020
	100	56	14	30
5 year	0.0050	0.0026	0.0015	0.0009
	100	52	29	19
5y5y	0.0039	0.0018	0.0023	-0.0002
	100	47	59	-6

the model can lead to quite large differences in the estimates. In our analysis, therefore, we use an average of four differently-specified models, two of which supplement the yield curve data with survey information. Specifically we use the benchmark and survey models in Malik and Meldrum (2016), the model in Vlieghe (2016); and the model in Andreasen and Meldrum (2015). Table 6 shows how much variation the two parts of the decomposition explain at each interest rate horizon and indicates that term premiums are relatively more important as the horizon of the interest rate increases.

We now repeat the exercise above on the change in expectations and term premiums seperately. First, we repeat the first stage regressions. Table 7 shows the first stage regressions for the decomposition of each asset. We also repeat the information test carried out in section **??** above. We find that the information test is significant at all maturities for both yield curve components.[10]

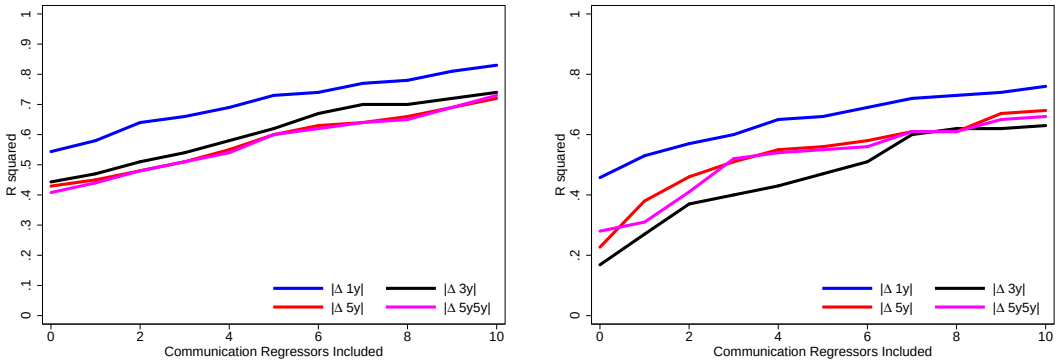
Second, we explore the important content for each asset class. Figure 11 replicates figure 8b from above showing, for the two elements of the decomposition, the R^2 as more topic characteristics are added. There are two key results. The quantitative signal, ω_t, affects expectations more strongly, and the narrative signal, χ_t. is especially important for longer-maturity term premiums.

Next we show using the Spearman rank correlations for each asset decomposition between overall features (as used above) and component-specific features. Table 8 shows the results. The results we found earlier for the overall asset response are generally reflecting expectations effects for shorter maturity assets and term premiums for longer-maturity assets. For the 5y5y yields, figure 12 shows the top three topics for each part of the decomposition.

[10]In the interest of space, we do not reproduce all the charts here.

Table 7: Post-Selection Topic Evaluation: 1 year spot

Main Regressors	(1) $\lvert Exp1y \rvert$	(2) $\lvert TP1y \rvert$	(3) $\lvert Exp3y \rvert$	(4) $\lvert TP3y \rvert$	(5) $\lvert Exp5y \rvert$	(6) $\lvert TP5y \rvert$	(7) $\lvert Exp5y5y \rvert$	(8) $\lvert TP5y5y \rvert$
π^f_{t+2}	-0.016	0.0018	-0.027	0.0045	-0.017	0.018	-0.021	0.031
$\lvert\Delta\pi_t^{MPC,+8} - \Delta\pi_t^{P,+8}\rvert$	0.040	-0.00034	-0.014	-0.019	-0.015	-0.020	-0.012	-0.019
$\text{Var}(\pi^f_{t+2})$	-0.0100	-0.0019	0.013	0.0084	0.012	0.0091	0.0063	0.012
$\Delta\text{Var}(\pi^f_{t+2})$	0.036	0.0027	0.0078	-0.035	-0.014	0.020	0.0070	0.046
$\text{Skew}(\pi^f_{t+2})$	-0.00092	0.0011	0.021	0.0016	0.012	0.013	0.0051	0.018
$\Delta\text{Skew}(\pi^f_{t+2})$	0.017	0.0077	0.035	0.0076	0.041	-0.0014	0.025	-0.00030
$\hat{y}^f_{t+2}$	0.028*	0.0012	0.012	-0.0079	0.0017	-0.0094	0.0014	-0.0075
$\lvert\Delta y_t^{MPC,+8} - \Delta y_t^{P,+8}\rvert$	-0.049**	-0.0025	0.018	0.019	0.027	0.0031	0.016	-0.017
$\text{Var}(\hat{y}^f_{t+2})$	0.0019	-0.00081	-0.0021	-0.0051	-0.00050	-0.0082	-0.00093	-0.013
$\Delta\text{Var}(\hat{y}^f_{t+2})$	0.0073	0.0021	0.0080	0.0067	0.0093	0.0029	0.0057	0.0083
$\text{Skew}(\hat{y}^f_{t+2})$	-0.0080	-0.0033	-0.039	-0.027	-0.049	-0.015	-0.034	-0.015
$\Delta\text{Skew}(\hat{y}^f_{t+2})$	-0.056*	-0.0060	-0.011	0.011	0.00064	-0.0016	0.00020	-0.019
$y, \tilde{+}8_t$	-0.034	0.19	2.44*	0.32	2.60**	-0.94	1.64	-1.93**
$\lvert y, \tilde{+}8_t - y, \tilde{+}8_{t-1}\rvert$	3.41*	0.99**	2.88	-0.79	2.13	0.0058	1.70	0.97
$\lvert y, \tilde{+}7_t - y, \tilde{+}8_{t-1}\rvert$	1.09	0.64	-0.29	0.15	-0.17	0.98	0.24	1.85
VIX	0.0010*	0.00030**	0.0014	0.00078	0.0011	0.00087	0.00096	0.00075
Constant	-0.060	-0.0019	-0.042	0.021	-0.019	0.035	-0.010	0.044
R-squared	0.544	0.457	0.443	0.168	0.429	0.227	0.408	0.280
R-squared No Vix	0.513	0.390	0.396	0.115	0.387	0.172	0.366	0.252



(a) Nominal Yields: Expectations **(b)** Nominal Yields: Term Premiums

Figure 11: Effect of adding more controls: R^2 - Nominal Yields Decomposition

7 Conclusion

Using the Bank of England Inflation Report, we show that there is there an information effect of central bank communication beyond conjunctural forecast information. While numerical forecasts drive changes across the yield curve, the information effects of forecast information are much less effect at the long end. Instead, there is an important role for narrative communication and, importantly, different dimensions of the narrative matter at different maturities. The information effect drives both longer-term expectations as well as term premiums; the latter suggests that communication likely also conveys distributional informational.

Our findings indicate the presence of a long-run information effect of central bank

Table 8: Spearman rank correlations between overall features and component-specific features

	Expectations	Term Premium
1yspot	0.943	0.463
3y	0.721	0.562
5y	0.451	0.325
5y5y	0.170	0.263

(a) Expectations 1: D7 **(b)** Expectations 2: D29 **(c)** Expectations 3: D27

(d) Term Premiums 1: D12 **(e)** Term Premiums 2: D20 **(f)** Term Premiums 3: D4

Figure 12: Post-Selection Topic Evaluation: 5y5y

text. This is supportive of open mouth operations and opens up the potential that central banks could engage in policy-free forward guidance. That is, they could talk down longer term interest rates without the need to confirm how plicy will react to macroeconomic conditions as is typically the case with forward guidance. For example, Carvalho et al. (2016) find that once US interest rates reached their zero-lower bound, communication continued to have effects on longer-maturity bonds even when shorter-maturity bonds stopped responding.

References

Andreasen, M. and Meldrum, A. (2015). Market beliefs about the uk monetary policy lift-off horizon: a no-arbitrage shadow rate term structure model approach. *Bank of England working papers*, 541.

Bansal, R. and Shaliastovich, I. (2013). A long-run risks explanation of predictability puzzles in bond and currency markets. *The Review of Financial Studies*, 26(1):1–33.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives*, 28(2):29–50.

Bernanke, B. S. and Kuttner, K. N. (2005). What explains the stock market's reaction to federal reserve policy? *Journal of Finance*, 60(3):1221–1257.

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

Blinder, A. S. (2008). Talking about Monetary Policy: The Virtues (and Vices?) of Central Bank Communication. Working Papers, Princeton University, Department of Economics, Center for Economic Policy Studies. 1048, Princeton University, Department of Economics, Center for Economic Policy Studies.

Blinder, A. S., Ehrmann, M., Fratzscher, M., Haan, J. D., and Jansen, D.-J. (2008). Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence. *Journal of Economic Literature, American Economic Association*, 46(4):910–45.

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

Campbell, J., Evans, C., Fisher, J., and Justiniano, A. (2012). Macroeconomic Effects of Federal Reserve Forward Guidance. *The Brookings Papers on Economic Activity*, Spring:1–54.

Carvalho, C., Hsu, E., and Nechio, F. (2016). Measuring the effect of the zero lower bound on monetary policy. Working Paper Series 2016-6, Federal Reserve Bank of San Francisco.

Chang, J., Gerrish, S., Boyd-Graber, J. L., and Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Advances in Neural Information Processing Systems*.

Cloyne, J. and Hürtgen, P. (2016). The macroeconomic effects of monetary policy: A new measure for the united kingdom. *American Economic Journal: Macroeconomics*, 8(4):75–102.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4):1047–1108.

Cook, T. and Hahn, T. (1989). The effect of changes in the federal funds rate target on market interest rates in the 1970s. *Journal of Monetary Economics*, 24(3):331–351.

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

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(Suppl. 1):5228–5235.

Gürkaynak, R. S., Sack, B., and Swanson, E. (2005). 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 McMahon, M. (2016). Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication. In *NBER International Seminar on Macroeconomics 2015*, NBER Chapters. National Bureau of Economic Research, Inc.

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

Hanson, S. G. and Stein, J. C. (2015). Monetary policy and long-term real rates. *Journal of Financial Economics*, 115(3):429–448.

Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. Number 143 in Monographs on Statistics and Applied Probability. CRC Press.

Jarociński and Karadi, P. (2017). Central Bank Information Shocks. Technical report, mimeo.

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

Laubach, T. and Williams, J. C. (2003). Measuring the natural rate of interest. *The Review of Economics and Statistics*, 85(4):1063–1070.

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

Malik, S. and Meldrum, A. (2016). Evaluating the robustness of uk term structure decompositions using linear regression methods. *Journal of Banking and Finance*, 67:85–102.

Martin, I. W. R. (2013). Consumption-based asset pricing with higher cumulants. *The Review of Economic Studies*, 80(2):745–773.

Meinshausen, N. and Bühlmann, P. (2006). High-Dimensional Graphs and Variable Selection with the LASSO. *The Annals of Statistics*, 34(3):1436–1462.

Meinshausen, N. and Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society Series B*, 72(4):417–473.

Melosi, L. (2017). Signalling Effects of Monetary Policy. *Review of Economic Studies*, 84(2):853–884.

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

Nakamura, E. and Steinsson, J. (2013). High Frequency Identification of Monetary Non-

Neutrality. NBER Working Papers 19260, National Bureau of Economic Research, Inc.

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

Reeves, R. and Sawicki, M. (2007). Do Financial Markets React to Bank of England Communication? *European Journal of Political Economy*, 23(1):207–227.

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

Romer, C. D. and Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B*, 58(1):267–288.

Vlieghe, G. (2016). Monetary policy expectations and long term interest rates.

Woodford, M. (2001). Monetary policy in the information economy. *Proceedings - Economic Policy Symposium - Jackson Hole*, pages 297–370.

Zhang, D. (2017). Term Structure, Forecast Revision and the Information Channel of Monetary Policy. Technical report, mimeo.

A Inflation Report Event Study

In this section, we conduct an event study to assess the average market impact of IR publication and other Bank of England communications. This extends the work of Reeves and Sawicki (2007), who conduct a similar analysis on a shorter sample. See section **??** in the main text for related discussion. We define the following events within our sample period: (1) IR publication; (2) policy rate announcement; (3) speech by MPC member; (4) release of minutes of MPC meeting. We define a dummy variable for each event, and estimate the model

$$|\Delta \text{Yield}|_t = \alpha + \beta_1 D(\text{IR})_t + \beta_2 D(\text{Rate})_t + \beta_3 D(\text{Speech})_t + \beta_4 D(\text{Min})_t + \varepsilon_t \tag{9}$$

for each yield. The estimated coefficients from ordinary least squares (OLS) estimates are in column (1) of tables 9a-10b. In columns (2)-(5) of these tables we estimate quantile regressions at various points in the distribution.

Confirming the visual evidence from the kernel densities in section **??**, at shorter maturities the IR is a dominant mover of market interest rates. The OLS coefficients for the one-year spot and three-year forward rates are both highly significant and approximately twice as large as the coefficient for policy announcements. There is a drop in significance for the five-year forward rate, but the magnitude of the IR coefficient is equivalent to that for policy announcements. This suggests a lack of power given there are three times as many announcements as IR dates over the sample period. However, there is a significant effect of IR releases in the right tail, as seen in column (5). For the five-year ahead, five year forward rate there are no signficiant coefficients, although the magnitude of the coefficient in column (5) is again the largest of any type of communication.

Table 9: Estimated Coefficients of Event-Study Regression

(a) 1-year spot rate

	(1)	(2)	(3)	(4)	(5)
Main Regressors	|Δ gilt1yspot|	|Δ gilt1yspot|	|Δ gilt1yspot|	|Δ gilt1yspot|	|Δ gilt1yspot|
IR	0.016***	0.0073	0.014***	0.019**	0.031
	[0.000]	[0.317]	[0.002]	[0.013]	[0.151]
Announcement	0.0084***	0.00076	0.0023	0.0033	0.038
	[0.002]	[0.522]	[0.236]	[0.394]	[0.149]
Speech	-0.0022**	-0.0013**	-0.0019**	-0.0024	-0.0034
	[0.033]	[0.039]	[0.030]	[0.108]	[0.396]
Minutes	0.0046**	-0.0011	0.0019	0.0065**	0.024***
	[0.029]	[0.302]	[0.344]	[0.035]	[0.002]
VIX	0.00098***	0.00027***	0.00063***	0.0013***	0.0030***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.0024*	0.0022***	0.0039***	0.0040**	0.0022
	[0.093]	[0.001]	[0.000]	[0.015]	[0.664]
R-squared	0.121				
Quantile	OLS	.25	.5	.75	.95
Sample	All	All	All	All	All

(b) 3-year forward rate

	(1)	(2)	(3)	(4)	(5)
Main Regressors	|Δ gilt3y|	|Δ gilt3y|	|Δ gilt3y|	|Δ gilt3y|	|Δ gilt3y|
IR	0.018***	0.0041	0.011	0.027**	0.061***
	[0.005]	[0.327]	[0.137]	[0.050]	[0.000]
Announcement	0.0071***	0.0027	0.011***	0.0097***	0.0088
	[0.007]	[0.274]	[0.000]	[0.004]	[0.339]
Speech	0.0039**	-0.0016	0.0025	0.0068***	0.022***
	[0.030]	[0.143]	[0.312]	[0.008]	[0.005]
Minutes	0.0042	0.0016	0.0028	0.012**	-0.0027
	[0.109]	[0.404]	[0.464]	[0.031]	[0.591]
VIX	0.00092***	0.00025***	0.00091***	0.0013***	0.0030***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.022***	0.0098***	0.014***	0.030***	0.045***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	0.053				
Quantile	OLS	.25	.5	.75	.95
Sample	All	All	All	All	All

Table 10: Estimated Coefficients of Event-Study Regression

(a) 5-year forward rate

	(1)	(2)	(3)	(4)	(5)
Main Regressors	|Δ gilt5y|	|Δ gilt5y|	|Δ gilt5y|	|Δ gilt5y|	|Δ gilt5y|
IR	0.0065	0.0019	-0.0039	0.0069	0.043***
	[0.271]	[0.578]	[0.449]	[0.557]	[0.002]
Announcement	0.0063*	0.0032	0.0059	0.010***	0.0032
	[0.055]	[0.233]	[0.106]	[0.004]	[0.857]
Speech	0.0044**	-0.00041	0.0030	0.0057**	0.025***
	[0.023]	[0.753]	[0.181]	[0.042]	[0.000]
Minutes	0.0037	0.0046	0.0050*	0.0040	-0.0055*
	[0.159]	[0.112]	[0.096]	[0.283]	[0.075]
VIX	0.0010***	0.00026***	0.00071***	0.0014***	0.0028***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.021***	0.0097***	0.019***	0.031***	0.052***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	0.052				
Quantile	OLS	.25	.5	.75	.95
Sample	All	All	All	All	All

(b) 5-year ahead, 5-year forward rate

	(1)	(2)	(3)	(4)	(5)
Main Regressors	|Δ gilt5y5y|	|Δ gilt5y5y|	|Δ gilt5y5y|	|Δ gilt5y5y|	|Δ gilt5y5y|
IR	0.0021	-0.0015	-0.0014	-0.0014	0.037*
	[0.688]	[0.592]	[0.798]	[0.889]	[0.074]
Announcement	0.0049	0.000084	0.0014	0.0078	0.0046
	[0.154]	[0.978]	[0.641]	[0.179]	[0.669]
Speech	0.0035*	-0.00050	3.7e-06	0.0026	0.021***
	[0.070]	[0.685]	[0.998]	[0.437]	[0.000]
Minutes	0.0036	0.0065*	0.0047*	0.0043	-0.0012
	[0.158]	[0.081]	[0.051]	[0.491]	[0.838]
VIX	0.00097***	0.00027***	0.00061***	0.0013***	0.0027***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.021***	0.0094***	0.020***	0.032***	0.051***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	0.049				
Quantile	OLS	.25	.5	.75	.95
Sample	All	All	All	All	All

B Quantitative Variables on TP and Expectations

Main Regressors	(1) $\lvert Exp1y \rvert$	(2) $\lvert TP1y \rvert$	(3) $\lvert Exp3y \rvert$	(4) $\lvert TP3y \rvert$	(5) $\lvert Exp5y \rvert$	(6) $\lvert TP5y \rvert$	(7) $\lvert Exp5y5y \rvert$	(8) $\lvert TP5y5y \rvert$
π^f_{t+2}	-0.016	0.0018	-0.027	0.0045	-0.017	0.018	-0.021	0.031
	[0.515]	[0.544]	[0.338]	[0.773]	[0.437]	[0.304]	[0.294]	[0.161]
$\lvert\Delta\pi_t^{MPC,+8} - \Delta\pi_t^{P,+8}\rvert$	0.040	-0.00034	-0.014	-0.019	-0.015	-0.020	-0.012	-0.019
	[0.162]	[0.950]	[0.669]	[0.280]	[0.568]	[0.256]	[0.617]	[0.348]
$\mathrm{Var}(\pi^f_{t+2})$	-0.0100	-0.0019	0.013	0.0084	0.012	0.0091	0.0063	0.012
	[0.313]	[0.203]	[0.305]	[0.241]	[0.242]	[0.262]	[0.458]	[0.222]
$\Delta\mathrm{Var}(\pi^f_{t+2})$	0.036	0.0027	0.0078	-0.035	-0.014	0.020	0.0070	0.046
	[0.400]	[0.780]	[0.851]	[0.153]	[0.676]	[0.624]	[0.791]	[0.397]
$\mathrm{Skew}(\pi^f_{t+2})$	-0.00092	0.0011	0.021	0.0016	0.012	0.013	0.0051	0.018
	[0.975]	[0.880]	[0.526]	[0.935]	[0.646]	[0.564]	[0.844]	[0.522]
$\Delta\mathrm{Skew}(\pi^f_{t+2})$	0.017	0.0077	0.035	0.0076	0.041	-0.0014	0.025	-0.00030
	[0.609]	[0.256]	[0.287]	[0.708]	[0.128]	[0.945]	[0.291]	[0.989]
$\hat{y}^f_{t+2}$	0.028*	0.0012	0.012	-0.0079	0.0017	-0.0094	0.0014	-0.0075
	[0.052]	[0.573]	[0.308]	[0.342]	[0.867]	[0.415]	[0.882]	[0.599]
$\lvert\Delta y_t^{MPC,+8} - \Delta y_t^{P,+8}\rvert$	-0.049**	-0.0025	0.018	0.019	0.027	0.0031	0.016	-0.017
	[0.040]	[0.618]	[0.455]	[0.177]	[0.206]	[0.870]	[0.384]	[0.469]
$\mathrm{Var}(\hat{y}^f_{t+2})$	0.0019	-0.00081	-0.0021	-0.0051	-0.00050	-0.0082	-0.00093	-0.013
	[0.860]	[0.690]	[0.867]	[0.482]	[0.962]	[0.325]	[0.915]	[0.186]
$\Delta\mathrm{Var}(\hat{y}^f_{t+2})$	0.0073	0.0021	0.0080	0.0067	0.0093	0.0029	0.0057	0.0083
	[0.608]	[0.267]	[0.589]	[0.339]	[0.335]	[0.715]	[0.548]	[0.439]
$\mathrm{Skew}(\hat{y}^f_{t+2})$	-0.0080	-0.0033	-0.039	-0.027	-0.049	-0.015	-0.034	-0.015
	[0.792]	[0.582]	[0.264]	[0.157]	[0.111]	[0.463]	[0.183]	[0.514]
$\Delta\mathrm{Skew}(\hat{y}^f_{t+2})$	-0.056*	-0.0060	-0.011	0.011	0.00064	-0.0016	0.00020	-0.019
	[0.053]	[0.344]	[0.696]	[0.526]	[0.979]	[0.928]	[0.992]	[0.393]
$y,\tilde{+}8_t$	-0.034	0.19	2.44*	0.32	2.60**	-0.94	1.64	-1.93**
	[0.975]	[0.335]	[0.080]	[0.696]	[0.025]	[0.219]	[0.105]	[0.040]
$\lvert y,\tilde{+}8_t - y,\tilde{+}8_{t-1}\rvert$	3.41*	0.99**	2.88	-0.79	2.13	0.0058	1.70	0.97
	[0.062]	[0.016]	[0.204]	[0.607]	[0.276]	[0.997]	[0.294]	[0.507]
$\lvert y,\tilde{+}7_t - y,\tilde{+}8_{t-1}\rvert$	1.09	0.64	-0.29	0.15	-0.17	0.98	0.24	1.85
	[0.555]	[0.105]	[0.903]	[0.924]	[0.933]	[0.548]	[0.890]	[0.327]
VIX	0.0010*	0.00030**	0.0014	0.00078	0.0011	0.00087	0.00096	0.00075
	[0.059]	[0.012]	[0.147]	[0.253]	[0.201]	[0.168]	[0.180]	[0.112]
Constant	-0.060	-0.0019	-0.042	0.021	-0.019	0.035	-0.010	0.044
	[0.113]	[0.746]	[0.311]	[0.478]	[0.612]	[0.323]	[0.752]	[0.266]
R-squared	0.544	0.457	0.443	0.168	0.429	0.227	0.408	0.280
R-squared No Vix	0.513	0.390	0.396	0.115	0.387	0.172	0.366	0.252

Table 11: First Stage Regressions by TP and Expectation

C Effect of Term Premium and Expectations Topics on Overall News

Main Regressors	(1) $\lvert\Delta 1yspot\rvert$	(2) $\lvert\Delta 1yspot\rvert$	(3) $\lvert\Delta 1yspot\rvert$	(4) $\lvert\Delta 1yspot\rvert$
T29		1.64**		1.32**
		[0.015]		[0.025]
d7		-0.57	-0.55	
		[0.214]	[0.289]	
T18		-1.30**		-1.17**
		[0.019]		[0.036]
T8		-1.75**		-2.56***
		[0.014]		[0.001]
d9		1.15		1.64**
		[0.198]		[0.039]
T2			-0.16	
			[0.747]	
d13			2.32	
			[0.150]	
d22			1.51	
			[0.179]	
d29			-1.15	
			[0.257]	
T13				-2.08***
				[0.010]
Constant	-0.070*	0.043	-0.085*	0.18**
	[0.075]	[0.507]	[0.053]	[0.019]
R-squared	0.563	0.706	0.618	0.740
Topics	-	Overall	Term Premia	Expectation
Include Levels of IR	Yes	Yes	Yes	Yes
Include Δ of IR	Yes	Yes	Yes	Yes
Partial R^2	.	0.34	0.13	0.40
Marg Pr Include	.	0.90	0.83	0.87

Table 12: 1yspot: Explanatory power of key topic features from Decompositions

Main Regressors	(1) $\lvert\Delta 3y\rvert$	(2) $\lvert\Delta 3y\rvert$	(3) $\lvert\Delta 3y\rvert$	(4) $\lvert\Delta 3y\rvert$
d7		-1.15		-1.29
		[0.253]		[0.163]
d29		-3.24*		-3.02*
		[0.071]		[0.072]
T9		1.71**	1.31	
		[0.034]	[0.124]	
d27		2.38*		2.58**
		[0.062]		[0.042]
d16		1.52		
		[0.112]		
d4			0.58	
			[0.539]	
T22			-0.42	
			[0.655]	
d20			-1.10	
			[0.640]	
T26			-0.55	
			[0.590]	
T29				1.76
				[0.131]
d15				3.57
				[0.108]
Constant	-0.026	-0.089	-0.025	-0.092
	[0.693]	[0.184]	[0.740]	[0.176]
R-squared	0.368	0.541	0.461	0.494
Topics	-	Overall	Term Premia	Expectation
Include Levels of IR	Yes	Yes	Yes	Yes
Include Δ of IR	Yes	Yes	Yes	Yes
Partial R^2	.	0.27	0.15	0.19
Marg Pr Include	.	0.76	0.77	0.84

Table 13: 3y: Explanatory power of key topic features from Decompositions

Main Regressors	(1) $\lvert\Delta 5y\rvert$	(2) $\lvert\Delta 5y\rvert$	(3) $\lvert\Delta 5y\rvert$	(4) $\lvert\Delta 5y\rvert$
T9		1.46**		1.51**
		[0.021]		[0.027]
d4		1.45*	1.07	
		[0.077]	[0.342]	
d20		-2.62	-2.50	
		[0.193]	[0.231]	
T6		1.49*		
		[0.058]		
d7		-0.85		-0.59
		[0.316]		[0.573]
d12			-2.96*	
			[0.053]	
d6			0.78	
			[0.515]	
T4			0.87	
			[0.403]	
d27				0.38
				[0.774]
d29				-1.93
				[0.266]
T14				-0.59
				[0.388]
Constant	0.014	-0.049	-0.028	-0.024
	[0.833]	[0.475]	[0.783]	[0.711]
R-squared	0.280	0.513	0.399	0.433
Topics	-	Overall	Term Premia	Expectation
Include Levels of IR	Yes	Yes	Yes	Yes
Include Δ of IR	Yes	Yes	Yes	Yes
Partial R^2	.	0.32	0.17	0.21
Marg Pr Include	.	0.66	0.81	0.76

Table 14: 5y: Explanatory power of key topic features from Decompositions

Main Regressors	(1) $\lvert\Delta 5y5y\rvert$	(2) $\lvert\Delta 5y5y\rvert$	(3) $\lvert\Delta 5y5y\rvert$	(4) $\lvert\Delta 5y5y\rvert$
d4		1.71**	1.49**	
		[0.013]	[0.042]	
d20		-3.23*	-3.25*	
		[0.064]	[0.094]	
T6		1.86***		
		[0.003]		
d12		-2.89**	-2.56*	
		[0.019]	[0.087]	
T9		0.72		
		[0.124]		
d10			0.97*	0.67
			[0.096]	[0.337]
T23			-1.58**	
			[0.043]	
d7				0.029
				[0.976]
d29				-2.17
				[0.166]
d27				0.25
				[0.826]
T29				0.33
				[0.741]
Constant	0.043	-0.017	0.072	0.0091
	[0.456]	[0.744]	[0.178]	[0.890]
R-squared	0.274	0.571	0.473	0.304
Topics	-	Overall	Term Premia	Expectation
Include Levels of IR	Yes	Yes	Yes	Yes
Include Δ of IR	Yes	Yes	Yes	Yes
Partial R^2	.	0.41	0.27	0.04
Marg Pr Include	.	0.69	0.85	0.84

Table 15: 5y5y: Explanatory power of key topic features from Decompositions

D Results Excluding the QE Period (Ending November 2008)

Main Regressors	(1) $\lvert\Delta 1y\rvert$	(2) $\lvert\Delta 3y\rvert$	(3) $\lvert\Delta 5y\rvert$	(4) $\lvert\Delta 5y5y\rvert$
π^f_{t+2}	-0.056*	-0.041	0.0013	0.0012
	[0.055]	[0.368]	[0.973]	[0.973]
$\lvert\Delta\pi_t^{MPC,+8} - \Delta\pi_t^{P,+8}\rvert$	-0.074*	-0.12**	-0.083*	-0.078
	[0.076]	[0.014]	[0.084]	[0.141]
$\mathrm{Var}(\pi^f_{t+2})$	-0.12**	-0.14	-0.073	-0.055
	[0.024]	[0.119]	[0.250]	[0.394]
$\Delta\mathrm{Var}(\pi^f_{t+2})$	0.070	0.041	-0.028	0.038
	[0.183]	[0.559]	[0.644]	[0.522]
$\mathrm{Skew}(\pi^f_{t+2})$	0.045	-0.057	-0.079*	-0.043
	[0.193]	[0.317]	[0.079]	[0.347]
$\Delta\mathrm{Skew}(\pi^f_{t+2})$	0.042	0.072**	0.072**	0.037
	[0.132]	[0.047]	[0.032]	[0.286]
$\hat{y}^f_{t+2}$	0.0099	-0.039*	-0.029	-0.023
	[0.486]	[0.087]	[0.178]	[0.246]
$\lvert\Delta y_t^{MPC,+8} - \Delta y_t^{P,+8}\rvert$	0.0040	0.036	0.036	0.019
	[0.861]	[0.454]	[0.414]	[0.660]
$\mathrm{Var}(\hat{y}^f_{t+2})$	0.11***	0.12***	0.067*	0.045
	[0.000]	[0.005]	[0.069]	[0.209]
$\Delta\mathrm{Var}(\hat{y}^f_{t+2})$	-0.067**	-0.072*	-0.034	-0.033
	[0.038]	[0.079]	[0.400]	[0.432]
$\mathrm{Skew}(\hat{y}^f_{t+2})$	-0.25***	-0.16*	-0.052	-0.063
	[0.000]	[0.062]	[0.442]	[0.411]
$\Delta\mathrm{Skew}(\hat{y}^f_{t+2})$	0.14***	0.049	-0.017	-0.020
	[0.005]	[0.491]	[0.801]	[0.781]
$y,\tilde{+}8_t$	0.74	-0.34	-1.54	-2.63
	[0.588]	[0.880]	[0.589]	[0.334]
$\lvert y,\tilde{+}8_t - y,\tilde{+}8_{t-1}\rvert$	13.6***	10.1***	3.65	1.58
	[0.000]	[0.003]	[0.186]	[0.547]
$\lvert y,\tilde{+}7_t - y,\tilde{+}8_{t-1}\rvert$	-6.19***	-3.99	-2.68	-1.94
	[0.003]	[0.271]	[0.442]	[0.544]
VIX	-0.00037	-0.00037	-6.0e-06	-2.4e-08
	[0.503]	[0.735]	[0.996]	[1.000]
Constant	-0.086**	0.053	0.063	0.071
	[0.029]	[0.381]	[0.283]	[0.212]
R-squared - Pre-QE	0.896	0.649	0.481	0.420
R-squared No Vix - Pre-QE	0.895	0.648	0.481	0.420
R-squared - All	0.563	0.368	0.280	0.274
R-squared No Vix - All	0.526	0.303	0.229	0.215

Table 16: First Stage Regressions - Excluding QE period

Main Regressors	(1) $\lvert Exp1y \rvert$	(2) $\lvert TP1y \rvert$	(3) $\lvert Exp3y \rvert$	(4) $\lvert TP3y \rvert$	(5) $\lvert Exp5y \rvert$	(6) $\lvert TP5y \rvert$	(7) $\lvert Exp5y5y \rvert$	(8) $\lvert TP5y5y \rvert$
π^f_{t+2}	-0.050*	-0.0014	-0.055*	0.019	-0.042	0.019	-0.037	0.012
	[0.080]	[0.856]	[0.095]	[0.301]	[0.153]	[0.420]	[0.137]	[0.711]
$\lvert\Delta\pi_t^{MPC,+8} - \Delta\pi_t^{P,+8}\rvert$	-0.051	-0.018**	-0.092**	-0.035	-0.067**	-0.047	-0.056*	-0.052
	[0.192]	[0.043]	[0.022]	[0.176]	[0.044]	[0.128]	[0.061]	[0.169]
$\mathrm{Var}(\pi^f_{t+2})$	-0.099**	-0.014	-0.11*	0.019	-0.094	0.037	-0.092*	0.024
	[0.035]	[0.357]	[0.098]	[0.515]	[0.103]	[0.356]	[0.076]	[0.683]
$\Delta\mathrm{Var}(\pi^f_{t+2})$	0.063	0.0074	0.073	-0.018	0.065	0.045	0.066	0.097
	[0.151]	[0.688]	[0.184]	[0.610]	[0.168]	[0.313]	[0.126]	[0.115]
$\mathrm{Skew}(\pi^f_{t+2})$	0.031	-0.0062	0.032	-0.033	0.0031	0.015	0.0076	0.054
	[0.364]	[0.653]	[0.455]	[0.287]	[0.933]	[0.648]	[0.834]	[0.172]
$\Delta\mathrm{Skew}(\pi^f_{t+2})$	0.033	0.013	0.022	0.0054	0.030	-0.032	0.016	-0.051**
	[0.245]	[0.175]	[0.463]	[0.815]	[0.198]	[0.147]	[0.473]	[0.043]
$\hat{y}^f_{t+2}$	0.0075	-0.0029	-0.022	-0.019*	-0.019	-0.019	-0.020	-0.015
	[0.597]	[0.495]	[0.234]	[0.092]	[0.199]	[0.130]	[0.162]	[0.329]
$\lvert\Delta y_t^{MPC,+8} - \Delta y_t^{P,+8}\rvert$	-0.0031	-0.0011	0.044	0.017	0.038	0.030	0.020	0.029
	[0.877]	[0.902]	[0.176]	[0.459]	[0.227]	[0.204]	[0.484]	[0.336]
$\mathrm{Var}(\hat{y}^f_{t+2})$	0.095***	0.013*	0.097***	0.0024	0.075***	-0.0057	0.069**	-0.0028
	[0.000]	[0.099]	[0.004]	[0.886]	[0.010]	[0.790]	[0.011]	[0.929]
$\Delta\mathrm{Var}(\hat{y}^f_{t+2})$	-0.062**	-0.0062	-0.067**	0.0049	-0.059**	-0.016	-0.041*	-0.037
	[0.049]	[0.542]	[0.020]	[0.841]	[0.026]	[0.573]	[0.071]	[0.342]
$\mathrm{Skew}(\hat{y}^f_{t+2})$	-0.21***	-0.027	-0.21***	-0.013	-0.15**	-0.062	-0.12**	-0.12*
	[0.000]	[0.139]	[0.005]	[0.765]	[0.023]	[0.198]	[0.034]	[0.073]
$\Delta\mathrm{Skew}(\hat{y}^f_{t+2})$	0.12***	0.014	0.11**	-0.0044	0.069	0.037	0.063	0.062
	[0.004]	[0.411]	[0.038]	[0.915]	[0.189]	[0.402]	[0.180]	[0.308]
$y,\tilde{+}8_t$	0.061	-0.076	0.92	-1.64	1.24	-2.02	0.51	-2.00
	[0.962]	[0.859]	[0.577]	[0.291]	[0.364]	[0.164]	[0.682]	[0.231]
$\lvert y,\tilde{+}8_t - y,\tilde{+}8_{t-1}\rvert$	10.5***	2.09***	11.0***	0.76	7.70***	2.89*	6.47***	4.00*
	[0.000]	[0.002]	[0.000]	[0.607]	[0.001]	[0.090]	[0.005]	[0.074]
$\lvert y,\tilde{+}7_t - y,\tilde{+}8_{t-1}\rvert$	-5.85***	0.31	-5.32**	-1.17	-2.86	-2.44	-2.17	-2.87
	[0.002]	[0.684]	[0.047]	[0.512]	[0.214]	[0.186]	[0.351]	[0.238]
VIX	-0.00044	8.0e-06	-0.00092	-0.00036	-0.00038	-0.00079	-0.00034	-0.00090
	[0.410]	[0.966]	[0.291]	[0.544]	[0.585]	[0.175]	[0.615]	[0.191]
Constant	-0.070*	0.0011	0.010	0.062*	0.015	0.067*	0.027	0.061
	[0.081]	[0.923]	[0.835]	[0.074]	[0.699]	[0.073]	[0.476]	[0.157]
R-squared - Pre-QE	0.890	0.518	0.773	0.260	0.682	0.455	0.653	0.525
R-squared No Vix - Pre-QE	0.888	0.518	0.763	0.251	0.679	0.426	0.650	0.504
R-squared - All	0.544	0.457	0.443	0.168	0.429	0.227	0.408	0.280
R-squared No Vix - All	0.513	0.390	0.396	0.115	0.387	0.172	0.366	0.252

Table 17: First Stage Regressions - Excluding QE period