Monetary policy and financial markets: evidence from Twitter traffic

Donato Masciandaro, Davide Romelli,
Gaia Rubera

TEP Working Paper No. 1023

June 2023

Trinity Economics Papers
Department of Economics

Monetary policy and financial markets: evidence from Twitter traffic*

Donato Masciandaro — Bocconi University

Davide Romelli† — Trinity College Dublin

Gaia Rubera — Bocconi University

Abstract

Monetary policy announcements of major central banks trigger substantial discussions about the policy on social media. In this paper, we use machine learning tools to identify Twitter messages related to monetary policy in a short-time window around the release of policy decisions of three major central banks, namely the ECB, the US Fed and the Bank of England. We then build an hourly measure of similarity between the tweets about monetary policy and the text of policy announcements that can be used to evaluate both the ex-ante predictability and the ex-post credibility of the announcement. We show that large differences in similarity are associated with a higher stock market and sovereign yield volatility, particularly around ECB press conferences. Our results also show a strong link between changes in similarity and asset price returns for the ECB, but less so for the Fed or the Bank of England.

JEL Classifications: E44, E52, E58, G14, G15, G41.

Keywords: monetary policy, central bank communication, financial markets, social media, Twitter, US Federal Reserve, European Central Bank, Bank of England.

* We would like to thank Marianna Blix Grimaldi, Michael Ehrmann, Laurent Ferrara, Etienne Farvaque, Paul Hubert, and Thomas Renault for stimulating discussions. We are grateful to the seminar participants at University College Dublin Smurfit Business School, University of Hagen, University of Lille, University of Nantes, University of Poitiers and Trinity College Dublin, as well as conference participants at the Computational and Financial Econometrics 2020 Virtual Conference, Western Economic Association International 2021 Virtual conference, Banque de France Real-Time Data Analysis, Methods and Applications Virtual Conference, 3rd Monash-Warwick-Zurich Text-as-Data Workshop, 4th Annual COMPTEXT Conference, 26th International Conference on Macroeconomic Analysis and International Finance, EEA-ESEM Congress 2022, 25th Central Bank Macroeconomic Modelling Workshop, the CEPR Central Bank Communication RPN panel 'Twitter and Central Banks' Communication' and the 2nd PSB Workshop on Banks and Financial Markets for helpful discussions, comments, and suggestions. Adriana Manna, Jacopo Maria Madurno and Romano Vincenzo Tarsia provided excellent research assistance. Davide Romelli gratefully acknowledges financial support from the 2020 Trinity College Dublin Arts and Social Sciences Benefactions Fund. Any errors or omissions are solely those of the authors.

†Corresponding author: Department of Economics, Trinity College Dublin, Arts Building, Dublin 2, Ireland. email: romellid@tcd.ie.

1 Introduction

Central banks' actions have never been more closely monitored than nowadays when social media enables monetary policy announcements to be widely communicated to the public. While central bank press releases are known to be closely monitored by financial market participants, the increasing use of social media by these participants can provide a useful setting to understand the effectiveness of monetary policy communication.

Two key features of monetary policy communication are the extent to which markets anticipate the policy (ex-ante predictability) and to what extent market expectations are aligned with the announcement after its release (ex-post credibility) (Svensson, 2014). In this paper, we propose a novel approach to assess these fundamental properties of monetary policy communication using Twitter activity around monetary policy decisions. Specifically, we use natural language processing techniques to construct measures of similarity between social media traffic around policy announcements and the text of the announcement. Our method aims to reveal the degree of consistency between Twitter messages and central bank policy statements as a way of assessing the predictability and credibility of monetary policy communication. We then investigate how discrepancies between the policy and market expectations are correlated with financial market behaviour around monetary policy announcements.

We collect all Twitter messages that contain keywords related to monetary policy in the four days around the policy announcements of three major central banks: the European Central Bank (ECB), the U.S. Federal Reserve Bank (Fed) and the Bank of England (BoE), between January 2011 and February 2020. We use machine learning tools to filter the initial sample of about half a million tweets to around 230,000 tweets that discuss topics related to the policy announcements of the three central banks in our sample. We then use a natural language processing algorithm to compute an hourly measure of the similarity between the text of a central bank announcement and our sample of tweets about that announcement. This measure of similarity can infer the alignment of Twitter users' expectations with the monetary policy decision before and after the policy announcement. We then interpret the time variation in text similarity as a proxy of Twitter users' perception of the degree of predictability and credibility of the policy announcement. Finally, we investigate whether social media users' reactions to central bank communication captured by our measure of similarity are aligned with financial market behaviour.

Our results show that changes in similarity are strongly correlated with financial market returns and volatility. In particular, announcements characterised by higher absolute changes in the measure of similarity before and after the announcement are associated with higher stock market variance and absolute returns, particularly following ECB press conferences. We also find that changes in tweets similarity are linked to higher stock market variance following announcements made by the Fed, but no effect is found for the Bank of England.

Our results also show a link between changes in similarity and sovereign bond yields. We find that changes in similarity around ECB press conferences are associated with larger changes in the realised variance and absolute returns of bond yields of four major Euro area countries, i.e. France, Germany, Italy and Spain. The magnitude of this correlation is stronger for longer-term sovereign bonds, suggesting that the market surprise captured by the changes in tweets similarity is more likely to be reflected in longer-maturity assets.

This work is related to a burgeoning literature that examines the impact of communication by central banks via Twitter. These recent works exploit high-frequency social media data to understand the effects of central bank communication by analyzing both the content created *by* monetary policy authorities, as well as the content by social media users *about* policy decisions. The bulk of evidence suggests that central banks have made significant progress in using social media to reach a wider public. This increased social media engagement, particularly around policy announcements, is shown to affect expectations, as well as behaviour in financial markets (Ehrmann and Hubert, 2022; Ehrmann and Wabitsch, 2022).

An important recent avenue of research in this literature relies on computational text analysis tools and machine learning techniques to analyse social media content. This literature provides strong evidence that social media data can reveal useful information about expectation formation and market sentiment around monetary policy events. Most evidence relies on high-frequency identification from social media traffic around major monetary policy events such as the U.S. "taper tantrum" or the ECB's Outright Monetary Transactions (OMT) announcement as well as other regular monetary policy announcements (Meinusch and Tillmann, 2017; Lüdering and Tillmann, 2020; Stiefel and Vivès, 2021; Ehrmann and Hubert, 2022; Ehrmann and Wabitsch, 2022). These methodologies highlight the usefulness of social media content in understanding how central bank communication via social media affects expectations and, subsequently, financial market outcomes.

The key contribution of our paper with respect to this existing research is the use of Twitter data to understand the extent to which discrepancies between market expectations, captured through social media discussions, and monetary policy decisions are associated with asset price volatility and returns. To the best of our knowledge, the only other paper that uses our methodological approach is Giavazzi et al. (2020), which computes measures of textual similarity between the tweets of German voters and the ones of the main German parties.

The remainder of this paper is structured as follows. A review of the related literature is provided in Section 2. Section 3 introduces our database of central bank communication events, discusses the Twitter data, presents the methodology used to construct the measure of similarity and describes the intraday data on equity and sovereign bonds. Section 4 presents the empirical findings, while Section 5 concludes.

2 Related literature

Our paper is related to the literature on the importance and effects of central bank communication. In the 1970s and 1980s, central banks were shrouded in monetary mystique and secrecy (Goodfriend, 1986). However, the development of modern monetary policy theory naturally produced a shift in communication from secrecy towards transparency (Eijffinger and Masciandaro, 2014) and central bank communication gained momentum (Blinder et al., 2008). Consequently, most central banks in advanced economies have taken major steps to incorporate communication strategies into their decision-making processes (Ehrmann and Fratzscher, 2005).

The increased importance of communication for policymakers is mirrored in the rapid development of the academic literature on this topic. This literature investigates the impact of central bank communication on macroeconomic variables, such as exchange rates (Jansen and De Haan, 2004; Fratzscher, 2008; Conrad and Lamla, 2010; Gürkaynak et al., 2021), interest rates (Gürkaynak, 2005; Gürkaynak et al., 2005; Lucca and Trebbi, 2009; Hayo and Neuenkirch, 2011; Lamla and Sturm, 2013; Neuenkirch, 2013; Altavilla et al., 2014; Lucca and Moench, 2015; Altavilla et al., 2019; Hansen et al., 2019), asset prices (Hayo et al., 2010; Rosa, 2011; Cieslak and Schrimpf, 2019; Ehrmann and Talmi, 2020; Gürkaynak et al., 2021; Gorodnichenko et al., 2023) and other real variables (Hansen and McMahon, 2016).

The literature has also analysed several aspects of central bank communication, such as consistency and tone of communication. For instance, Jansen and de Haan (2013) analyse whether the ECB uses consistent language in its communication and find an overall consistency, even though it seems flexible enough to adapt to changing circumstances. Acosta and Meade (2015) study the similarity of FOMC post-meeting statements and show that they have become more similar over time, especially since the global financial crisis. Nevertheless, FOMC statements have also become more complex since the onset of unconventional monetary policies, as shown by Hernández-Murillo et al. (2014). More recently, language processing algorithms have been used to identify differences between subsequent FOMC statements (Doh et al., 2022). Similarly, computational linguistic tools have been used to analyse the tone of monetary policy communication (Bailey and Schonhardt-Bailey, 2008; Lucca and Trebbi, 2009; Schonhardt-Bailey, 2013; Gerlach, 2004; Kawamura et al., 2016; Hansen and McMahon, 2016; Hansen et al., 2018; Schmeling and Wagner, 2019; Hubert and Labondance, 2021; Bailliu et al., 2021; Gáti and Handlan, 2022).

More recent literature looks at social media interactions and monetary policy communication (see Masciandaro et al., 2023, for a review). A part of this literature has focused on the use of social media as a complementary communication channel *by* central banks. Korhonen and Newby (2019) examine the extent to which European central banks maintain an institutional Twitter account and analyse their tweeting activity. They find that central banks' Twitter activity has no relation to citizens' online participation and that communication on financial stability has increased more in comparison to the one on monetary policy. Looking at the United States, Conti-Brown and Feinstein (2020) undertake the first systematic analysis of the Fed's participation on Twitter and find that the Fed is more engaged on Twitter than other independent agencies. Gorodnichenko et al. (2021) analyse the Federal Reserve System communication on Facebook and Twitter and its effectiveness. In the case of the Fed, communication via Twitter appears to be more popular and gains greater public engagement. They show that market participants update their inflation expectations based on information contained in the Fed's social media posts. However, they find no evidence of stock market reactions to the Fed's communication on social media.

A second body of work focuses on tweets *about* monetary policy made by Twitter users. For example, Azar and Lo (2016) create a new dataset of tweets that cite the Fed to understand how investors on social media behave around FOMC meeting dates. Their results suggest that tweets contain information that can be used to predict returns and build portfolios that outperform the benchmark market portfolio. Meinusch and Tillmann (2017) and Lüdering and Tillmann (2020) analyse the Fed's *taper tantrum* period between April and October 2013 and capture information on the debate among market professionals during this period. They show that both the revisions of expectations of market participants as well as shocks to selected topics discussed in the tweets can lead to significant changes in U.S. bond yields, exchange rates, and stock prices. Similarly, Stiefel and Vivès (2021) study the extent to which changes in beliefs about ECB's OMT policy during the summer of 2012 can explain the sudden reduction in government bond spreads for distressed countries in the euro area. Ehrmann and Wabitsch (2022) analyse tweets about the ECB to understand the extent to which its communication is received by non-experts and how it affects their views. They show that Twitter communications by the ECB spark significant ECB-related traffic, which tends to be more factual and less subjective. Ehrmann and Hubert (2022) use the database created by Ehrmann and Wabitsch (2022) to investigate how ECB-related tweets made in the days preceding an ECB press conference are associated with the magnitude of the monetary policy surprises on the announcement day. In particular, using data on disagreement about the economic outlook, they find that Twitter traffic is correlated with the size of monetary policy surprises as Twitter users pay more attention to meetings in which they expect larger changes in the monetary policy stance. Adams et al. (2023) use natural language processing tools on Twitter data to develop a new index to gauge financial market sentiment. They find that this index correlates with corporate bond spreads and other measures of financial conditions, and helps predict next-day stock market returns as well as future changes in the U.S. monetary policy stance. Finally, Renault et al. (2023) propose a new measure of intraday investor attention by analyzing Twitter messages related to ECB announcements. Their results suggest that when investor attention is high prior to the announcements, asset prices experience greater absolute changes. Additionally, they show that Twitter is a more reliable source for measuring attention compared to other sources.

To the best of our knowledge, our paper is the first to investigate how discrepancies in market beliefs based on social media interactions and the policy announcement are related to asset price volatility and returns in the hours surrounding monetary policy announcements.

3 Similarity of tweets and monetary policy announcements: methodology and data

This section describes the steps followed in the construction of the sample of Twitter messages about monetary policy that we employ to compute our measure of similarity between market beliefs and central bank communication. We also present descriptive statistics on the high-frequency data on stock market indices and government bond prices.

3.1 Monetary policy communication

We first create a database of time-stamped communication on monetary policy decisions by three central banks: the European Central Bank, the U.S. Federal Reserve Bank and the Bank of England. Our sample period runs from January 2011 through February 2020.[1]

The first part of our paper investigates the variation in tweets' similarity around scheduled monetary policy announcements. We focus exclusively on scheduled events as these events normally attract significant social media traffic both before and after the announcement, while this is not the case for unscheduled monetary policy announcements. Our database includes (i) 89 press releases published by the ECB at 13:45 (Frankfurt time) and 89 transcripts of the press conference that begins at 14:30 and ends at 15:30 (Frankfurt time); (ii) 71 press releases issued by the Fed following each FOMC meeting at 14:00 (New York time), and (iii)) 94 monetary policy decisions made by the BoE at 12:00 (London time).[2]

[1]The decision to start our analysis in 2011 is motivated by the limited availability of tweets before 2011, while we choose to stop our analysis in February 2020 to exclude the extraordinary measures taken by central banks since the start of the Covid-19 pandemic.

[2]Since the FOMC meeting of March 19-20, 2013, FOMC statements are released at 14:00 New York time. Before this date, the Fed press releases were published at either 12:30 or 14:15. The exact timing of each press release has been taken into consideration for the extraction of the associated tweets. The

3.2 Tweets on monetary policy

We use the GetOldTweets Python package to collect all English-language tweets related to monetary policy published in the 48-hour window around the scheduled monetary policy announcements of the three central banks of interest between January 2011 and February 2020. We first manually analyse a random sample of ten monetary policy announcements in order to identify keywords and hashtags used by Twitter users to discuss monetary policy decisions. Next, we automatise the selection of tweets by collecting all Twitter messages that: (a) mentioned the official Twitter account of the central bank, e.g. @bankofengland; (b) contained a hashtag followed by the central bank's acronym, e.g. #ecb; or (c) contained a hashtag followed by the surname of the chair of the central bank, e.g. #yellen.[3] Table 1 presents an overview of the keywords used to extract tweets.The overall number of tweets collected during this first round of the selection process is 467,777.

Table 1: Overview of the keywords used for Twitter messages extraction

Central bank	Keywords		
European Central Bank	@ecb	#ecb	#trichet #draghi #lagarde
Federal Reserve	@federalreserve	#fed	#bernanke, #yellen, #powell
Bank of England	@bankofengland	#boe, #bankofengland	#carney

3.2.1 Selection of relevant tweets

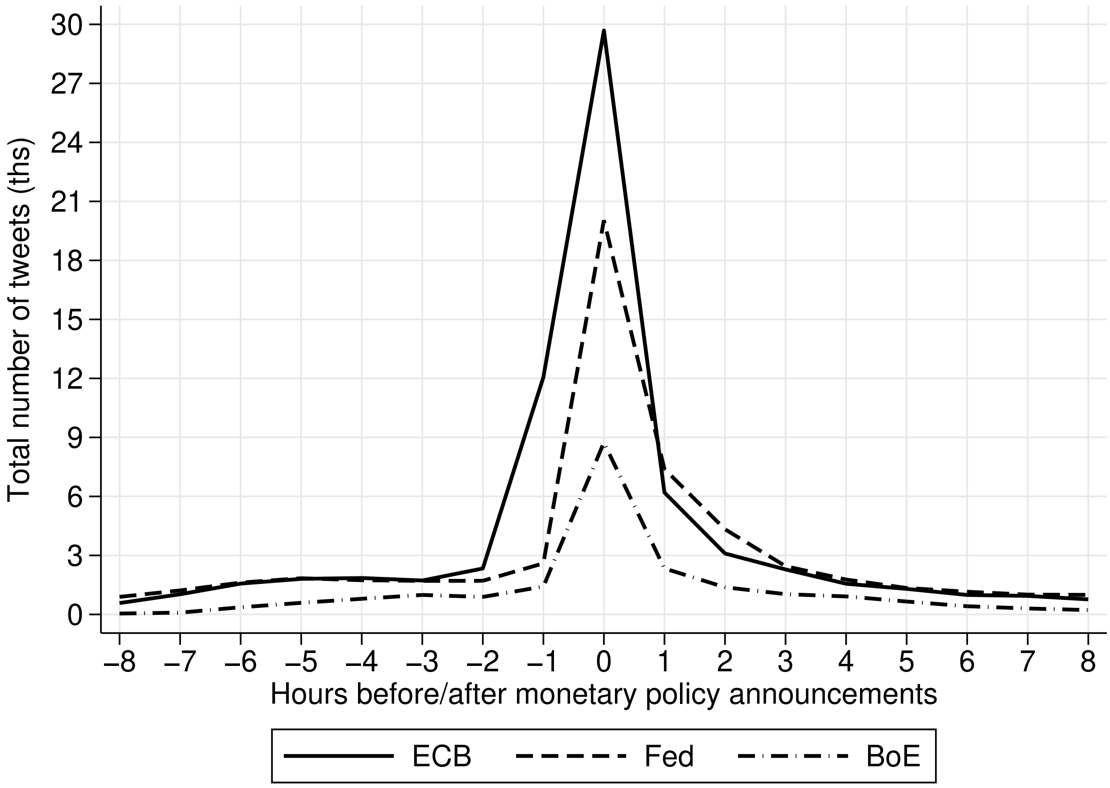
As not all of the collected tweets were related to the policy announcement, we trained a machine learning algorithm on a manually labelled training set to isolate a subset of relevant tweets. To do so, we first selected a random sample of 3,000 tweets and we asked two research assistants to independently classify tweets as relevant, i.e. related to monetary policy announcements, or irrelevant. Details on the guiding principles used for the selection of the relevant tweets are presented in Appendix A. At the end of the classification process, we considered as relevant the following set of tweets: 1) considered relevant by both the research assistants and the authors of the paper, or 2) classified as relevant by one of the two research assistants, and validated by the authors. This screening process allowed us to identify 782 relevant, i.e. 26% of the sub-sample and 2,218 irrelevant tweets.

timing of these events has been double-checked with the data in Cieslak and Schrimpf (2019).

[3] Given the high number of tweets potentially associated with the surname of the former Governor of the Bank of England, Mervyn King (#King), we decided to exclude this hashtag from the search.

After manually labelling the sub-sample of tweets, we used the classification to train an algorithm to identify all relevant tweets within the entire corpus of 467,777 messages. In doing so, we identified 228,348 tweets that discuss the monetary policy decisions of the three central banks in our sample around each of their scheduled announcements between January 2011 and February 2020.[4] Figure 1 shows the total number of monetary policy-related tweets published in the eight hours around monetary policy announcements. Not surprisingly, the number of tweets spikes in the 2-hour window surrounding an announcement and this provides a first check that the supervised classifier employed to train the algorithm achieves a clear identification of relevant tweets.

Figure 1: Twitter traffic around central bank announcements



Note: The figure shows the total number of monetary policy-related tweets published in the 8-hour window around a central bank scheduled communication. The solid line refers to the tweets related to the ECB, the dashed line to those mentioning the Fed, while the dashed-dotted line refers to the BoE.

[4]The accuracy of the classification or the performance of the classification model, computed as the number of correct predictions divided by the total number of predictions, is 79%.

3.3 Tweets similarity

We use the relevant tweets identified to compute an hourly measure of similarity between these Twitter messages and the text of the monetary policy announcements by transforming the two corpora of text into vectors using Doc2Vec, a deep learning algorithm. Details on the pre-processing processes and the technique used to compute the measure of similarity are reported in Appendix B. Here, we briefly summarize the method.

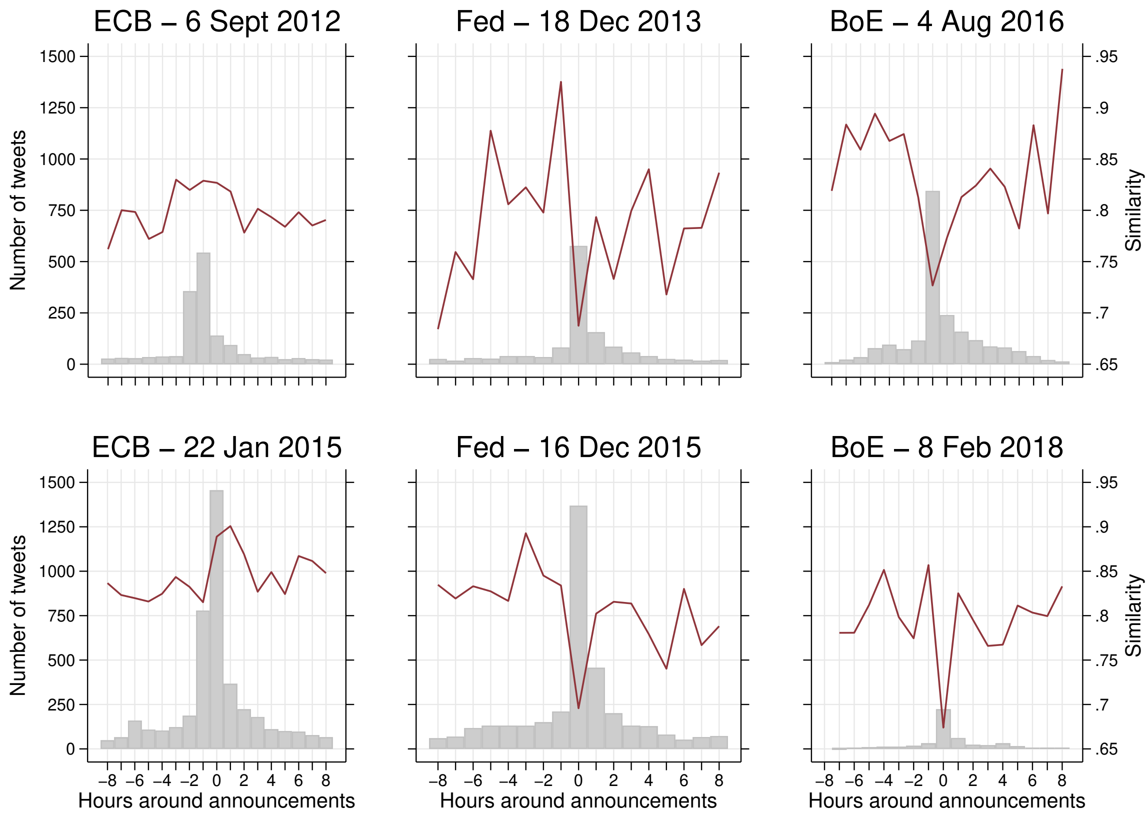
For each hour surrounding a monetary policy announcement, we create documents on: i) central bank transcripts and ii) tweets. Transcripts contain the text of the monetary policy decision released at a specific date and time by one of the three central banks in our sample, while the tweet documents aggregate the text of all the tweets related to monetary policy published in a given hour around an announcement. We use Doc2Vec an unsupervised deep-learning algorithm that learns how to represent each document as a unique vector (Le and Mikolov, 2014). We then measure similarity between documents as the cosine of the angle between the two corresponding vectors, i.e. the normalized inner product of the vector of the text of monetary policy communication c and the one of tweets on monetary policy t at hour h of day d:

$$\cos\theta_{tc_{d,h}} = \frac{\overrightarrow{t_{d,h}} \quad \overrightarrow{c_{d,h}}}{\left|\overrightarrow{t_{d,h}}\right| \left|\overrightarrow{c_{d,h}}\right|}. \tag{1}$$

We test the validity of our similarity measure by manually cross-checking six representative announcements (two for each central bank) in our sample. These events are: 1) the launch of the OMT programme by the ECB on September 6, 2012; 2) ECB's adoption of the expanded asset purchase programme on January 22, 2015;3) the Fed announcement to taper its bond-buying program on December 18, 2013; 4) the Fed's decision to raise the Federal funds rate for the first time since 2006 on December 16, 2015; 5) the BoE announcement to cut rates for the first time since 2009 on August 4, 2016; and 6) the BoE warning of the possibility of earlier and larger rate hikes on February 8, 2018. The evolution of the similarity measure around these events is reported in Figure 2.

In line with anecdotal evidence, the two announcements made by the ECB had been largely anticipated by the market participants and the general public. The low volatility of the similarity measure in the hours surrounding these events supports this evidence. In

Figure 2: Similarity measure validation: key events



Note: The figure shows the evolution of the similarity measure in the 8-hour window around six selected monetary policy decisions. The grey bars indicate the number of relevant tweets published in the hour of reference.

addition, we can notice that the similarity measure in the hour of the announcement (h=0) is higher for the ECB than for the events reported for both the Fed and the BoE. Indeed, the four selected announcements for these latter central banks had been less anticipated by the markets as also suggested by the spike in the number of tweets in the hour of the announcement, and this is reflected in a higher variation of the similarity measure around monetary policy events.

3.4 High-frequency data and asset price volatility and returns

Our empirical analysis aims at understanding the link between changes in similarity and asset market performance. To do so, we extract high-frequency, one-minute data on stock market indices and government bond prices from Refinitiv. The data availability and the coverage of maturities for government bonds differ from country to country. For the euro area, we have data on stock market indices for France, Germany, Italy and Spain as well as the EURO STOXX50 and the EURO STOXX Banks Index, which is the stock market index for the biggest banks in the euro area. We also obtained sovereign yields with maturities ranging from 1 to 30 years for these four major Euro area countries. For the US, we have the stock market indices for the Dow Jones, Nasdaq and S&P 500 and Treasury yields with maturities of 2, 5, 10 and 30 years, while for the UK we have high-frequency data for the FTSE 100 stock market index and Gilts yields with maturities of 1, 2, 5, 10, 15 and 30 years.[5]

Since our goal is to assess the sensitivity of asset prices to central bank communication, we first compute the realised variance of stock returns and bond yields around monetary policy announcements. Let τ denote the time of a communication event, and $\tau^- = \tau - h^-$ and $\tau^+ = \tau + h^+$ the time window before and after the event. If we divide the interval $h^+ + h^-$ into N sub-intervals of length $\Delta = \frac{h^+ + h^-}{N}$, then the Realized Variance (RV) of asset prices around event τ is computed as:

$$\mathrm{RV}_\tau\left(\tau^-, \tau^+, N\right) = \sum_{i=0}^{N} r^2_{\tau+i\Delta}, \tag{2}$$

where $r^2_{\tau+i\Delta} = \left(p_{\tau+i\Delta} - p_{\tau+(i-1)\Delta}\right)^2$ and p is the log of the asset price. In our baseline estimations we construct the realized variance by summing up the squared value of the one-minute returns over an event window: from 15 min before to 15 min after a monetary policy decision (for example between 13:30 and 14:00 for the European Central Bank press releases).[6]

An alternative way to assess the sensitivity of asset prices to monetary policy announcements is to look at their returns. Given that our measure of changes in similarity does not capture information on tweet sentiment, we focus our attention on the absolute change in prices. Specifically, we compute the absolute value of returns following Altavilla et al. (2019) who measure returns as the percentage variation in the median price between the 15-25 minutes following a press release and the 10-20 minutes prior to it. Following their

[5] Appendix Table C.1 provides information on the full set of data available in our database, together with their RIC (Reuters Identification Code) code.

[6] For the ECB press conference, we compute the realised variance between 14:15, i.e. 15 minutes before the conference, and 15.45, i.e. 15 minutes after the end of the event.

approach, returns associated with, for example, ECB press conferences are computed using the median price in the 14:15-14:25 interval as the pre-conference window and the median price in the 15:40-15:50 interval as the post-conference window.

4 Monetary Policy, tweets and asset price volatility and returns

In this section, we investigate the association between changes in tweets similarity and asset price volatility and returns in the hours surrounding monetary policy announcements.

4.1 Asset price volatility

We start by presenting the high-frequency identification strategy that exploits the link between changes in similarity and asset price variance. The estimation takes the following form:

$$\mathrm{RV}_{\tau,c} = \alpha_y + \beta_1 |\Delta \mathrm{Similarity}_{h,c}| + \epsilon_{\tau,c}, \tag{3}$$

where $RV_{\tau,c}$ is the realised variance of returns around the time τ of a communication event c, which in our baseline model is -15 and +15 minutes around a monetary policy announcement. The main explanatory variable is the absolute change in the measure of similarity between the hour post and prior to a monetary policy announcement, i.e. computed as $|\Delta Similarity_{h,c}| = |Similarity_{h,c}| - |Similarity_{h-1,c}|$. We interpret this absolute change in similarity prior to the event (ex-ante predictability) as compared to after the announcement (ex-post credibility) as a proxy of the extent of financial market surprise or disagreement with the policy. As such, we expect a higher absolute change in similarity to be associated with a stronger asset price volatility. We include in Eq. (3) year fixed-effects, α_y, to absorb common time-variation in asset price reactions to monetary policy announcements within a year and to control for the trend in Twitter usage. To avoid assigning excessive weight to monetary policy events that attracted limited social media traffic, we use a weighted least squares approach, weighting each event by the number of tweets in the hours surrounding a monetary policy announcement.

4.1.1 Stock market variance

We start by estimating the association between the realized variance of stock market indices and the absolute change in tweet similarity in the hour post and the hour prior to the release of a monetary policy decision. The realized variance of stock market indices is computed over a 30-minutes window around each event. Table 2 shows the results for the realized variance of the stock market indices of the four biggest economies of the Euro area, i.e. France, Germany, Italy, and Spain, and two Euro area stock indices for blue chip companies (STOXX50E) and banks (SX7E) in a 30-minute window around ECB announcements.

As discussed in Altavilla et al. (2019), the ECB policy decisions are announced in two separate steps. At 13:45 Central European Time (CET) a brief press release summarizes the policy decision without providing any explanation and rationale for the decision. Then, at 14:30 CET the ECB President reads the introductory statement, which explains the rationale behind the decision. Usually, the introductory statement is read out in about 15 minutes and the conference continues with a follow-up question-and-answer session of the ECB President with journalists that lasts for about 45 minutes. Until December 2014, press releases only provided information related to policy rate decisions, disregarding announcements on non-standard measures. Between January 2015 and January 2016 press releases mentioned the adoption of further measures but did not provide details, which were announced during the press conference. Finally, starting from March 2016, the content of the decisions on non-standard policy measures is also summarized in the press release, but all the details are provided during the introductory statement to the press conference. This staggered procedure motivates our decision to provide two estimates for the results related to the ECB: one for the press release window (panel A) and one for the press conference window (Panel B).

The results reported in Panel A of Table 2 suggest that the absolute change in similarity around press releases is not associated with the volatility in European stock market indices in the 30-minute window around this type of announcement. Consistent with the idea that more information is provided during the ECB press conference, the results presented in Panel B show a positive and strongly significant coefficient for absolute changes in similarity across all estimations. This suggests that larger changes in tweets' similarity before and after ECB press conferences are associated with higher stock market variance. These results

Table 2: Changes in similarity and Euro area stock market indices variance

Panel A: Press release window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	0.879	0.215	3.660	2.677	1.919	3.538
	(3.597)	(2.840)	(3.050)	(4.365)	(3.482)	(5.150)
Observations	89	89	89	89	89	89
R-squared	0.267	0.257	0.377	0.230	0.312	0.402

Panel B: Press conference window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	5.855***	5.023***	7.098***	6.643***	6.814***	11.941***
	(1.671)	(1.382)	(2.253)	(2.175)	(1.880)	(3.809)
Observations	89	89	89	89	89	89
R-squared	0.462	0.468	0.486	0.389	0.478	0.402

Note: The dependent variable is the realized variance of the stock market indices of major Euro area countries: CAC 40 for France, DAX for Germany, FTSE MIB for Italy, and IBEX for Spain, as well as the EURO STOXX50 (STOXX50E) and EURO STOXX Banks (SX7E) indices for European blue chip companies and banks in the 30-minute window around each event. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour prior to the release of a monetary policy decision. Year fixed-effects dummies are included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

support the idea that press conferences and in particular, Q&A sessions facilitate market participants' information processing and are associated with higher trading activity (Hayo et al., 2020).

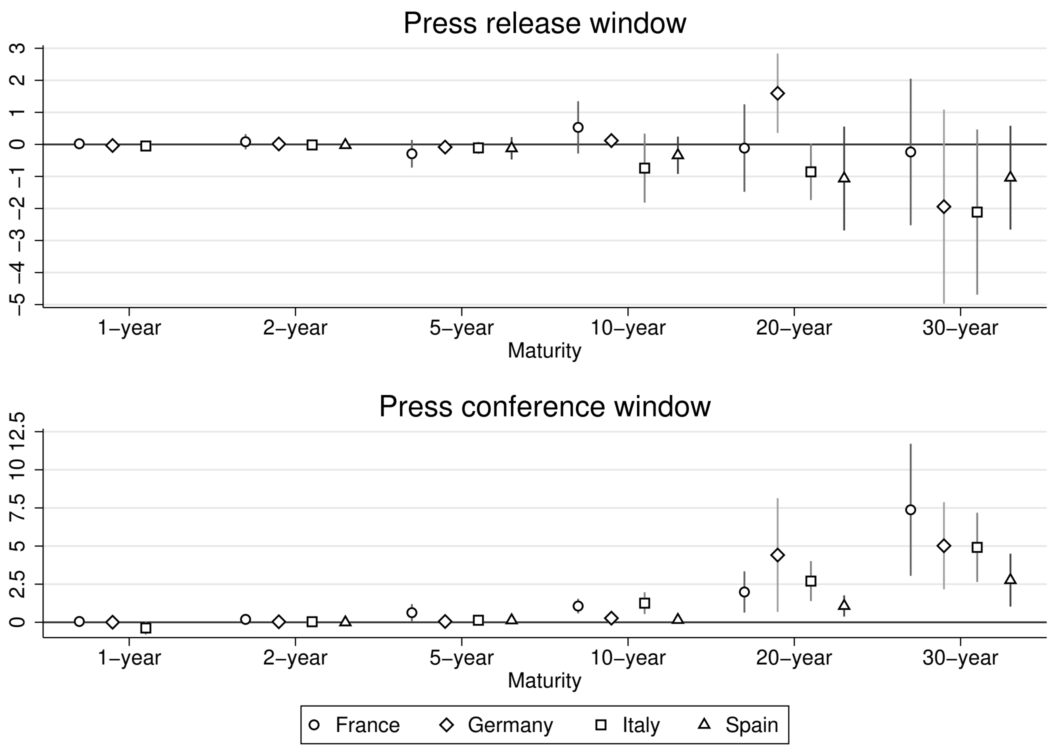
Table 3: Changes in similarity and US and UK stock market indices variance

	United States			United Kingdom
	(1) Dow Jones	(2) Nasdaq	(3) S&P 500	(4) FTSE 100
$\lvert\Delta$ Similarity$\rvert$	1.512**	1.181**	1.453**	0.060
	(0.609)	(0.550)	(0.583)	(0.075)
Observations	71	71	71	94
R-squared	0.589	0.622	0.593	0.240

Note: The dependent variable is the realized variance of US and UK stock market indices in the 30-minute window surrounding a monetary policy announcement. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour prior to the release of a monetary policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 3 reports the estimates for the realized variance of US and UK stock market indices. The results reported in Columns (1)-(3) suggest that larger changes in similarity are associated with a higher realized volatility of the US stock market indices following the publication of Federal Reserve press releases. The absolute change in tweets' similarity around the monetary policy announcements made by the Bank of England, on the other hand, does not seem to be associated with the realized variance of the FTSE 100 Index.

Figure 3: Changes in similarity and European sovereign yields volatility during ECB announcements



Note: The figure show the coefficient of |Δ Similarity| in Eq. (3). The dependent variable is the realized variance of European sovereign yields around the ECB press release window [13:30–14:00] and ECB press conference window [14:15–15:45], respectively. Year fixed-effects are included. 90% confidence intervals are presented.

4.1.2 Sovereign bond yields variance

Next, we estimate the association between changes in similarity and the realized variance of sovereign yields around monetary policy announcements. Figure 3 summarises the results for French, German, Italian and Spanish sovereign yields at different maturities around press releases (panel a) and press conferences (panel b) windows.[7] This figure shows some

[7]See Appendix Tables D.1-D.4 for information on the estimations obtained for each country at different maturities.

interesting patterns. Similar to the results shown in Table 2, Figure 3 documents that absolute changes in the measure of similarity are not associated with the realized variance of European sovereign bonds in the 30-minute window surrounding ECB press releases (panel a). The only exceptions are represented by the realised variance of the German sovereign yields with maturities of 1 and 20 years, respectively. On the other hand, the results presented in panel b of Figure 3 highlight a positive and statistically significant association between absolute changes in similarity and sovereign yield volatility at longer maturities, i.e. from 5 to 30 years. Importantly, the magnitude of the coefficient of interest, i.e. the absolute change in similarity, increases in magnitude for sovereign bonds characterized by longer maturities. This evidence suggests that the Twitter traffic surprise captured by the change in the measure of similarity is associated with higher volatility of long-term government bonds.

Table 4: Changes in similarity and US and UK sovereign yield volatility

Panel A: United States

	(1)	(2)	(3)	(4)
	2-year	5-year	10-year	30-year
$\lvert\Delta$ Similarity$\rvert$	-0.002	0.010	0.017	0.102
	(0.001)	(0.007)	(0.018)	(0.081)
Observations	67	71	71	71
R-squared	0.158	0.339	0.245	0.254

Panel B: United Kingdom

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	15-year	30-year
$\lvert\Delta$ Similarity$\rvert$	0.005	0.004	0.002	0.053	0.125	0.187
	(0.023)	(0.011)	(0.040)	(0.128)	(0.305)	(1.067)
Observations	94	94	94	9	94	94
R-squared	0.407	0.140	0.312	0.223	0.197	0.419

Note: The dependent variable is the realized variance of United States and United Kingdom sovereign yields at different maturities in the 30-minute window surrounding a monetary policy announcement. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour prior to the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

The results for the sovereign yield variance of US and UK government bonds are reported in Panel A and Panel B of Table 4, respectively. Consistent with the results presented in Figure 3, the magnitude of the coefficients increases at longer maturities. However, none of the estimated coefficients is statistically different from zero, suggesting a less strong link between market disagreement and the realized volatility of both US and UK government bond yields.

Overall, the estimates presented in this section highlight that larger changes in the similarity of tweets related to monetary policy are associated with higher bond yield volatility following ECB press conferences and suggest that the information content of monetary policy communication is key in driving the realized variance of these assets around monetary policy announcements. These results are in line with those found in previous literature, which suggest that the questions answered during the Q&A session have substantial effects on markets (Ehrmann and Fratzscher, 2007). As a matter of fact, during ECB press conferences the President of the ECB does not only provide information on the decisions taken during the Governing Council meetings (introductory statement), but they also answer questions asked by the attending journalists during the questions-and-answers (Q&A) session.

4.2 Asset price returns

The results presented so far focused on asset price variance. In this section, we explore the link between changes in the measure of similarity and asset price returns. The estimation takes the following form:

$$|\mathrm{r}|_{\tau,c} = \alpha_y + \beta_1 |\Delta \mathrm{Similarity}_{h,c}| + \epsilon_{\tau,c}; \tag{4}$$

where $|r|_{\tau,c}$ is the absolute value of returns obtained by computing the absolute percentage variation in the median price between the 15-25 minutes following a press release and the 10-20 minutes before it.[8] Similar to the estimations presented in section 4.1, we also add year fixed effects, α_y, and use a weighted least squares approach, weighting each event by the number of tweets in the hours surrounding a monetary policy announcement.

[8] This event window is similar to the one used in Altavilla et al. (2019). As both stock market indices and sovereign yields can experience positive or negative returns following monetary policy announcements, we focus our analysis on the absolute value of returns as our measure of tweets similarity does not capture information on the direction of monetary policy decisions or changes in the sentiment of Twitter messages, but only how close tweets are related to monetary policy announcements.

Table 5: Changes in similarity on Euro area stock market indices returns

Panel A: Press release window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	3.404	1.708	5.924	5.713*	4.184	10.557*
	(2.740)	(2.086)	(3.913)	(3.174)	(3.001)	(5.472)
Observations	89	89	89	89	89	89
R-squared	0.331	0.331	0.458	0.437	0.377	0.573

Panel B: Press conference window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	4.319**	4.941***	4.055**	2.297	2.418**	1.987
	(1.741)	(1.600)	(1.617)	(1.460)	(1.187)	(1.943)
Observations	89	89	89	89	89	89
R-squared	0.289	0.317	0.218	0.223	0.207	0.252

Note: The dependent variable is the absolute value of returns of the stock market indices of major Euro area countries, i.e. CAC 40 for France, DAX for Germany, FTSE MIB for Italy and IBEX for Spain, as well as the EURO STOXX50 and EURO STOXX Banks indices for the Euro area using high-frequency one-minute data. Returns are computed as the percentage variation in the median price between the 15-25 minutes following a press release and the 10-20 minutes before it. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 6: US and UK stock market indices returns

	United States			**United Kingdom**
	(1)	(2)	(3)	(4)
	Dow Jones	Nasdaq	S&P 500	FTSE 100
$\lvert\Delta$ Similarity$\rvert$	0.525	0.725	1.077	0.414
	(0.922)	(0.902)	(1.012)	(0.428)
Observations	71	71	71	94
R-squared	0.365	0.280	0.268	0.205

Note: The dependent variable is the absolute value of returns of US and UK stock market indices using high-frequency one-minute data. Returns are computed as the percentage variation in the median price between the 15-25 minutes following a press release and the 10-20 minutes before it. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

4.2.1 Stock market returns

Table 5 reports the estimations on the association between the absolute change in tweet similarity and the absolute returns of European stock market indices using high-frequency data. Similar to the results presented in Table 2, which suggested that large changes in similarity were associated with higher stock market variance during ECB press conferences, the results presented in Table 5 show a statistically significant relationship between changes in similarity and European stock market returns. The only exceptions are represented by the Spanish stock market index and the Stoxx index for Banks for which the estimated coefficient is statistically associated with stock market returns during press release windows, but not around press conferences.

Furthermore, the results presented in Table 6 show the absence of any link between changes in tweets similarity and stock market returns for both the United States and the United Kingdom. Overall, our results suggest a strong association between stock returns and our measure of changes in market expectations only following announcements made by the European Central Bank.

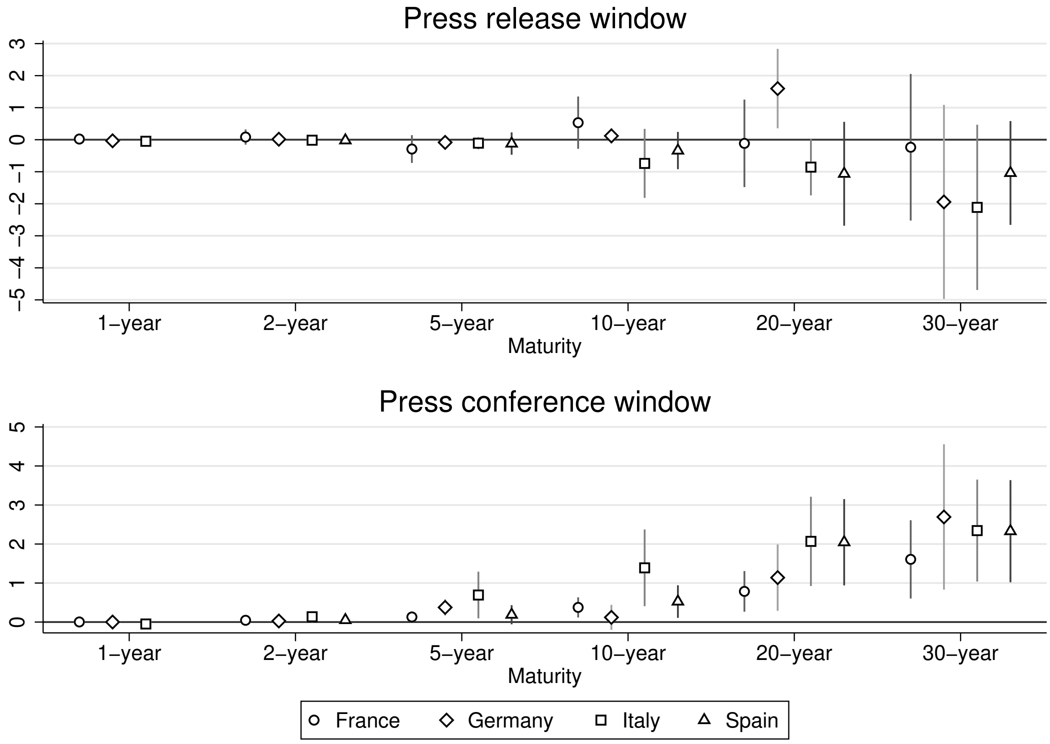
4.2.2 Sovereign bond returns

We now turn to investigate the link between sovereign bond returns and changes in tweets similarity.[9] Figure 4 shows the coefficients of changes in similarity estimated on the absolute returns of French, German, Italian and Spanish bonds at different maturities. The results presented in the top figure suggest that changes in tweet similarity are not associated with sovereign bond returns around ECB press releases. However, when we focus our attention on the press conference window (bottom figure), we observe a positive relationship between changes in similarity and the absolute change in bond prices. This effect is statistically significant mainly for sovereign bonds characterized by a maturity longer than 2 years.

Finally, Table 7 presents the estimates for the absolute change in yields for the United States and the United Kingdom. These results are in line with those in the previous section and show no statistical relationship between changes in similarity and sovereign bond returns for the Fed and the Bank of England.

[9] See Appendix Tables D.5-D.8 for information on the estimations obtained for each country at different maturities.

Figure 4: Changes in similarity and European sovereign yield returns during ECB announcements



Note: The figure show the coefficient of |Δ Similarity| in Eq. (3). The dependent variable is the absolute return of European sovereign yield around the ECB press release and press conference windows, respectively. Following Altavilla et al. (2019), sovereign yield returns are computed as the percentage variation in the median price between the 15-25 minutes following an announcement and the 10-20 minutes before it. Year fixed-effects are included. 90% confidence intervals are presented.

4.3 Robustness checks

To assess the sensitivity of our results, we perform a series of robustness checks. First, given the importance of defining the most appropriate event window for the analysis, the first category of robustness tests focuses on alternative event windows. Although our baseline estimations use event windows common in the literature, i.e. -15 to +15 min around an announcement, we test the robustness of our results by focusing on two different event windows: -10 to +10 min and -15 to +30 min around an announcement, respectively.[10] The results presented in Tables D.9 and D.10 and Figures D.1 and D.2 show that using

[10] For brevity, we only report the results obtained using the -15 to +30 min event window and focusing on the association between changes in similarity and asset price volatility and returns around European Central Bank announcements. The results for the Federal Reserve Bank and the Bank of England are unchanged using this alternative event window, as well as the -10 to +10 min one. These results are available upon request.

Table 7: Changes in similarity and US and UK sovereign bond returns

Panel A: United States

	(1) 2-year	(2) 5-year	(3) 10-year	(4) 30-year
ΔSimilarity	-0.029	0.042	0.046	-0.229
	(0.018)	(0.059)	(0.121)	(0.257)
Observations	67	71	71	71
R-squared	0.165	0.180	0.166	0.119

Panel B: United Kingdom

	(1) 1-year	(2) 2-year	(3) 5-year	(4) 10-year	(5) 15-year	(6) 30-year
|Δ Similarity|	-0.035	-0.041	0.088	0.289	0.331	-0.859
	(0.047)	(0.133)	(0.396)	(0.786)	(0.898)	(1.355)
Observations	91	91	91	91	91	91
R-squared	0.502	0.393	0.348	0.329	0.332	0.444

Note: The dependent variable is the absolute change in United States and United Kingdom sovereign bond returns at different maturities using high-frequency one-minute data. Returns are computed as the percentage variation in the median price between the 15-25 minutes following a press release and the 10-20 minutes before it. |Δ Similarity| is the change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies have been included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

alternative event windows does not affect our conclusion that larger changes in similarity are associated with higher asset prices volatility and absolute returns.

Finally, we verify that our results are not driven by higher market volatility on the day of the monetary policy announcement by controlling for stock markets volatility using the VSTOXX volatility index for the ECB and the Bank of England and the VIX for the Federal Reserve Bank. The results are also robust to the inclusion of this control variable (see Tables D.11 and D.12 and Figures D.3 and D.4).

5 Conclusion

Traditionally, central banks have mainly communicated with financial market participants that had a well-defined profile: a professional interest in following monetary policy information and the necessary knowledge to understand central bank messages (Ehrmann and Wabitsch, 2022). However, the widespread use of unconventional monetary policy tools in recent years has called for better explanations of what central banks do and has led many central banks to step up their efforts to communicate to a wider audience. As a result, central banks are increasingly engaging in social media as a regular feature of their communication policy.

At the same time, social media users have increased their use of these channels to express their opinions on various topics, including monetary policy decisions. In this paper, we propose a novel approach to investigating the reactions of social media users to central bank announcements and testing the association with asset price variance and returns using high-frequency data. Specifically, we employ machine learning techniques to compute a measure of textual similarity between tweets related to monetary policy and the text of monetary policy announcements from three major central banks: the European Central Bank, the Federal Reserve Bank, and the Bank of England.

Our results suggest that large changes in the similarity between tweets and monetary policy decisions before and after announcements are associated with higher asset price variance and returns, especially following announcements made by the European Central Bank during its press conferences. These findings are consistent with those found in previous literature, which suggests that the questions answered during the Q&A session have substantial effects on markets (Ehrmann and Fratzscher, 2007), and this might be related to the additional information provided during the questions-and-answers session.

The novel data and empirical strategy in this paper also highlight the usefulness of social media reactions of market participants as a proxy for the degree of market surprise or disagreement with the policy, and ultimately as a tool to assess the effectiveness of central bank communication.

References

Acosta, M. and E. E. Meade (2015). Hanging on every word: Semantic analysis of the fomc's post-meeting statement. Technical report, Board of Governors of the Federal Reserve System (US).

Adams, T., A. Ajello, D. Silva, and F. Vazquez-Grande (2023). More than words: Twitter chatter and financial market sentiment.

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

Altavilla, C., D. Giannone, and M. Lenza (2014). The financial and macroeconomic effects of omt announcements.

Azar, P. D. and A. W. Lo (2016). The wisdom of twitter crowds: Predicting stock market reactions to fomc meetings via twitter feeds. *The Journal of Portfolio Management 42*(5), 123–134.

Bailey, A. and C. Schonhardt-Bailey (2008). Does deliberation matter in fomc monetary policymaking? the volcker revolution of 1979. *Political Analysis*, 404–427.

Bailliu, J., X. Han, B. Sadaba, M. Kruger, et al. (2021). Chinese monetary policy and text analytics: Connecting words and deeds. Technical report, Bank of Canada.

Bhatia, S., J. H. Lau, and T. Baldwin (2016). Automatic labelling of topics with neural embeddings. *arXiv preprint arXiv:1612.05340*.

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

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

Conrad, C. and M. J. Lamla (2010). The high-frequency response of the eur-usd exchange rate to ECB communication. *Journal of Money, Credit and Banking 42*(7), 1391–1417.

Conti-Brown, P. and B. D. Feinstein (2020). Twitter and the federal reserve. *Brookings Center on Regulation & Markets Working Paper*.

Doh, T., D. Song, and S.-K. Yang (2022). Deciphering federal reserve communication via text analysis of alternative fomc statements. *Federal Reserve Bank of Kansas City Working Paper Forthcoming*.

Ehrmann, M. and M. Fratzscher (2005). Central bank communication: different strategies, same effectiveness. *ECB Working*.

Ehrmann, M. and M. Fratzscher (2007). Explaining monetary policy in press conferences.

Ehrmann, M. and P. Hubert (2022). Information acquisiion ahead of monetary policy announcements. Technical report, mimeo.

Ehrmann, M. and J. Talmi (2020). Starting from a blank page? semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics 111*, 48–62.

Ehrmann, M. and A. Wabitsch (2022). Central bank communication with non-experts – a road to nowhere? *Journal of Monetary Economics*.

Eijffinger, S. and D. Masciandaro (2014). *Modern Monetary Policy and Central Bank Governance*. Edward Elgar Publishing.

Fratzscher, M. (2008). Communication and exchange rate policy. *Journal of Macroeconomics 30*(4), 1651–1672.

Gáti, L. and A. Handlan (2022). Monetary communication rules. *Available at SSRN 4260246*.

Gerlach, S. (2004). Interest rate setting by the ECB: words and deeds.

Giavazzi, F., F. Iglhaut, G. Lemoli, and G. Rubera (2020). Terrorist attacks, cultural incidents and the vote for radical parties: Analyzing text from twitter. Technical report, National Bureau of Economic Research.

Goodfriend, M. (1986). Monetary mystique: Secrecy and central banking. *Journal of Monetary Economics 17*(1), 63–92.

Gorodnichenko, Y., T. Pham, and O. Talavera (2023). The voice of monetary policy. *American Economic Review 113*(2), 548–584.

Gorodnichenko, Y., T. Pham, O. Talavera, et al. (2021). 'liked', 'shared', 'commented': Central bank communication on facebook and twitter. Technical report.

Gürkaynak, R. S. (2005). Using federal funds futures contracts for monetary policy analysis.

Gürkaynak, R. S., A. H. Kara, B. Kısacıkoğlu, and S. S. Lee (2021). Monetary policy surprises and exchange rate behavior. *Journal of International Economics 130*, 103443.

Gürkaynak, R. S., B. P. Sack, and E. T. Swanson (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking 1*, 55–93.

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

Hansen, S., M. McMahon, and A. Prat (2018). Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics 133*(2), 801–870.

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

Hayo, B., K. Henseler, M. S. Rapp, and J. Zahner (2020). Complexity of ecb communication and financial market trading.

Hayo, B., A. M. Kutan, and M. Neuenkirch (2010). The impact of us central bank communication on european and pacific equity markets. *Economics Letters 108*(2), 172–174.

Hayo, B. and M. Neuenkirch (2011). Canadian interest rate setting: The information content of canadian and us central bank communication. *Southern Economic Journal 78*(1), 131–148.

Hernández-Murillo, R., H. Shell, et al. (2014). The rising complexity of the fomc statement. *Economic Synopses* (23).

Hubert, P. and F. Labondance (2021). The signaling effects of central bank tone. *European Economic Review 133*, 103684.

Jansen, D.-J. and J. De Haan (2004). Look who's talking: ECB communication during the first years of emu. *De Nederlandsche Bank, manuscript*.

Jansen, D.-J. and J. de Haan (2013). An assessment of the consistency of ECB communication. *Central Bank Communication, Decision Making, and Governance: Issues, Challenges, and Case Studies*, 183.

Kawamura, K., Y. Kobashi, M. Shizume, and K. Ueda (2016). Strategic central bank communication: discourse and game-theoretic analyses of the bank of japan's monthly report.

Korhonen, I. and E. Newby (2019). Mastering central bank communication challenges via twitter. Technical report, BoF Economics Review.

Lamla, M. J. and J.-E. Sturm (2013). Interest rate expectations in the media and central bank communication. *Central Bank Communication, Decision Making, and Governance: Issues, Challenges, and Case Studies*, 101–111.

Le, Q. and T. Mikolov (2014). Distributed representations of sentences and documents. In *International conference on machine learning*, pp. 1188–1196. PMLR.

Lucca, D. O. and E. Moench (2015). The pre-fomc announcement drift. *The Journal of Finance 70*(1), 329–371.

Lucca, D. O. and F. Trebbi (2009). Measuring central bank communication: an automated approach with application to fomc statements. Technical report, National Bureau of Economic Research.

Lüdering, J. and P. Tillmann (2020). Monetary policy on twitter and asset prices: Evidence from computational text analysis. *The North American Journal of Economics and Finance 51*, 100875.

Masciandaro, D., O. Peia, and D. Romelli (2023). Central bank communication and social media: From silence to twitter. *Journal of Economic Surveys*.

Meinusch, A. and P. Tillmann (2017). Quantitative easing and tapering uncertainty: Evidence from Twitter. *International Journal of Central Banking 13*(4), 227–258.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Neuenkirch, M. (2013). Monetary policy transmission in vector autoregressions: A new approach using central bank communication. *Journal of Banking & Finance 37*(11), 4278–4285.

Renault, T., M. Picault, and R. L. Gillet (2023). Investor attention and intraday market reaction to ecb announcements. *Available at SSRN 4415462*.

Rosa, C. (2011). Words that shake traders: The stock market's reaction to central bank communication in real time. *Journal of Empirical Finance 18*(5), 915–934.

Schmeling, M. and C. Wagner (2019). Does central bank tone move asset prices? *Available at SSRN 2629978*.

Schonhardt-Bailey, C. (2013). *Deliberating American monetary policy: A textual analysis.* MIT Press.

Stiefel, M. and R. Vivès (2021). "whatever it takes" to change belief: evidence from Twitter. *Review of World Economics*, 1–33.

Svensson, L. E. (2014). Forward guidance. Technical report, National Bureau of Economic Research.

Appendices

A Criteria for tweets manual classification

This section provides a summary of the guiding principles used for the classification of the 3,000 tweets which have been manually labelled as relevant or non-relevant, i.e. not related to monetary policy.

- **Conservativeness**: We focused on those tweets discussing either monetary policy or financial sector supervision (or both). However, we adopted a conservative approach by selecting only those tweets that were pertinent. For a tweet to be pertinent, it had to contain a description or a judgment over the course of action (expected or announced) by one of the central banks in our sample.

- **Machine thinking**: Given that our work is supposed to be replicated in an automated way on a large scale of tweets, we performed our assessment accordingly. Here follow the main considerations:

  1. We decided to classify as relevant only those tweets which were self-explanatory. As the machine would operate on a tweet-by-tweet basis (sort of row by row), reading exclusively the text of the message. Therefore, we decided to:
     (a) exclude those tweets which were part of a larger Twitter thread and were difficult to understand in isolation. Such types of tweets were mainly answers to a previous tweet or to a chain of tweets and their meaning was clear only after reading the entire thread.
     (b) exclude those tweets which included an image or an URL address and which required the image or the URL to be used in order to be properly understood.
  2. We classified as non-relevant all the tweets whose language style was excessively difficult to grasp due to their metaphorical phrasing, abbreviations, use of slang language, etc.

- **Accuracy and netiquette**: we deliberately excluded those tweets which were excessively generic.

- **Advertising, updates and market trends**: some of the tweets in our sample had advertising goals. Also, there were tweets whose objective was to update traders about the latest news and market trends. We classified both categories of tweets as non-relevant.

- **Papers, conferences and other policy documents**: we classified as non-relevant all tweets referring to conferences, papers and seminars.

B Text processing details

As discussed in the main text, we compute the similarity between the tweets related to monetary policy and the text of the monetary policy announcements by transforming the two corpora of text into vectors using Doc2Vec, a deep learning algorithm. As we are interested in how changes in similarity are associated with asset price variations around monetary policy announcements, we gathered all the English tweets published in the interval going from 48 hours before an announcement and the 48 hours following it. The tweets were then split into 1 hour segments around the monetary policy communication events.

We then measure similarity as the cosine similarity between each one-hour corpus of tweets and the nearest monetary policy decision. Before computing the measure of similarity using the Doc2Vec deep learning algorithm, we pre-processed our corpus of text. In this section, we provide information on the procedures followed.

B.1 Text pre-processing

With pre-processing we reduce the number of words, and hence the computational time necessary to run the Doc2Vec deep learning algorithm, without losing relevant information. We follow standard procedures in text pre-processing with different libraries in Python. First, we pre-processed the text of both central bank announcements and tweets by lowercasing all words. For tweets, we also removed all URLs and mentions to other Twitter users. We then transformed the text into single words called "tokens". Thereafter, we eliminated stop words, i.e. words that occur frequently in our corpus such as "and" and "the" but have little meaning and punctuations. We do this by using the "word_tokenize" module in the NLTK Python package. We also removed all tokens that consisted only of non-alphanumeric characters, emoticons as well as the @ and # symbols.

Next, we lemmatized words using the WordNetLemmatizer module in the NLTK Python package. Lemmatization entails reducing words to a common root form, called a "lemma", to limit the presence of synonyms. Then we performed stemming, which implies conflating the various forms of a word into a common representation known as the stem. For instance, as a result of this process, the words "ate" and "eating" are both reduced to the common stem "eat". Stemming and lemmatization rely on pre-existing dictionaries for the English language, which explains why we only collected English tweets using the GetOldTweets Python package. We relied on Porter stemming algorithm in the NLTK Python package for our stemming. Finally, we introduced collocation, i.e. the combination of two words that have higher probabilities of co-occurring together than separately, using the BigramCollocationFinder module in the NLTK Python package. For instance, the tokens "federal" and "reserve" have higher chances of co-occurring as the bigram "federal reserve" than appearing separately. In this case, collocations transform the two separate tokens into just one: "federal_reserve". We then use the pre-processed corpus of text to train the Doc2Vec deep learning algorithm.

B.2 Vector representation: Doc2Vec

After pre-processing our tweets and transcripts, we obtained two types of "documents": 1) the transcripts of the monetary policy decisions, i.e. the text of monetary policy announcements, and 2) the full set of one-hour tweets, i.e. the text of all the tweets published on a certain day-hour window around the central bank communication of reference.

Following Giavazzi et al. (2020), our approach consists of using neural networks to compute vector representations of words, including their context, through embedding. To perform this task, Mikolov et al. (2013) propose using Word2Vec, which learns word embeddings and aims to predict the occurrence of a word given the surrounding words (context). In this model, every word is mapped to a unique vector, which is represented by a column in the weight matrix W. The algorithm constructs a vocabulary from the input corpus and then learns word representations by training a neural network language model. The model is trained using stochastic gradient descent with backpropagation. When the model converges, it represents words as word embeddings, i.e. meaningful real-value vectors of configurable dimensions (usually 150-500 dimensions). The neural network learns a word's embedding based on its contexts in different sentences. As a result, the words that occur in similar contexts are mapped onto close vectors.

As an extension of the Word2Vec neural network, Le and Mikolov (2014) introduced Doc2Vec deep learning algorithm to learn embeddings of sentences and documents (or sentence embeddings), not just words. By treating each document as a word token, the Doc2Vec methodology is used to learn document embeddings (Bhatia et al., 2016). As with Word2Vec, training occurs through backpropagation. Each iteration of the algorithm is called an "epoch", and its purpose is to increase the quality of the output vectors. This type of document embedding allows for texts to be represented as dense, fixed-length feature vectors that take their semantic and syntactic structure into account. We used a freely available implementation of the Doc2Vec algorithm included in the GENSIM Python module and asked for 300-dimensional vectors.

C High-frequency data

Table C.1: High-frequency data available in the dataset

Central Bank	Country	RIC	Asset type
European Central Bank	France	FR10YT	Government bond
European Central Bank	France	FR15YT	Government bond
European Central Bank	France	FR1YT	Government bond
European Central Bank	France	FR20YT	Government bond
European Central Bank	France	FR2YT	Government bond
European Central Bank	France	FR30YT	Government bond
European Central Bank	France	FR3MT	Government bond
European Central Bank	France	FR5YT	Government bond
European Central Bank	France	FR6MT	Government bond
European Central Bank	France	.FCHI	Stock index
European Central Bank	Germany	DE10YT	Government bond
European Central Bank	Germany	DE15YT	Government bond
European Central Bank	Germany	DE1YT	Government bond
European Central Bank	Germany	DE20YT	Government bond
European Central Bank	Germany	DE2YT	Government bond
European Central Bank	Germany	DE30YT	Government bond
European Central Bank	Germany	DE3MT	Government bond
European Central Bank	Germany	DE5YT	Government bond
European Central Bank	Germany	DE6MT	Government bond
European Central Bank	Germany	.GDAXI	Stock index
European Central Bank	Italy	IT10YT	Government bond
European Central Bank	Italy	IT15YT	Government bond
European Central Bank	Italy	IT1YT	Government bond
European Central Bank	Italy	IT20YT	Government bond
European Central Bank	Italy	IT2YT	Government bond
European Central Bank	Italy	IT30YT	Government bond
European Central Bank	Italy	IT3MT	Government bond
European Central Bank	Italy	IT5YT	Government bond
European Central Bank	Italy	IT6MT	Government bond
European Central Bank	Italy	.FTMIB	Stock index
European Central Bank	Spain	ES10YT	Government bond
European Central Bank	Spain	ES20YT	Government bond
European Central Bank	Spain	ES2YT	Government bond
European Central Bank	Spain	ES30YT	Government bond
European Central Bank	Spain	ES3MT	Government bond
European Central Bank	Spain	ES5YT	Government bond
European Central Bank	Spain	ES6MT	Government bond
European Central Bank	Spain	.IBEX	Stock index

Note: The table reports information on the high-frequency data available in our dataset. RIC refers to the Reuters Identification Code used for the extraction of the data from Refinitiv.

Table C.1 continued: High-frequency data available in the dataset

Central Bank	**Country**	**RIC**	**Asset type**
Federal Reserve	United States	US10YT	Government bond
Federal Reserve	United States	US2YT	Government bond
Federal Reserve	United States	US30YT	Government bond
Federal Reserve	United States	US3MT	Government bond
Federal Reserve	United States	US5YT	Government bond
Federal Reserve	United States	US6MT	Government bond
Federal Reserve	United States	.DJI	Stock index
Federal Reserve	United States	.INX	Stock index
Federal Reserve	United States	.IXIC	Stock index
Bank of England	United Kingdom	GB10YT	Government bond
Bank of England	United Kingdom	GB15YT	Government bond
Bank of England	United Kingdom	GB1YT	Government bond
Bank of England	United Kingdom	GB20YT	Government bond
Bank of England	United Kingdom	GB2YT	Government bond
Bank of England	United Kingdom	GB30YT	Government bond
Bank of England	United Kingdom	GB5YT	Government bond
Bank of England	United Kingdom	.FTSE	Stock index

Note: The table reports information on the high-frequency data available in our dataset. RIC refers to the Reuters Identification Code used for the extraction of the data from Refinitiv.

D Appendix Tables

Table D.1: Changes in similarity and French sovereign yield variance

Panel A: Press release window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\|\Delta$ Similarity$\|$	0.020	0.083	-0.292	0.531	-0.115	-0.236
	(0.023)	(0.143)	(0.260)	(0.490)	(0.821)	(1.375)
Observations	88	89	89	89	89	89
R-squared	0.417	0.412	0.329	0.384	0.429	0.586

Panel B: Press conference window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\|\Delta$ Similarity$\|$	0.052	0.190	0.631*	1.066***	1.989**	7.378***
	(0.078)	(0.161)	(0.338)	(0.287)	(0.812)	(2.601)
Observations	88	89	89	89	89	89
R-squared	0.125	0.494	0.364	0.464	0.378	0.552

Note: The dependent variable is the realised variance of French sovereign yields at different maturities in the 30-minute window surrounding a monetary policy announcement. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $|\Delta$ Similarity$|$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.2: Changes in similarity and German sovereign yield variance

Panel A: Press release window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\lvert\Delta$ Similarity$\rvert$	-0.034***	0.014	-0.083	0.120	1.596**	-1.942
	(0.012)	(0.014)	(0.078)	(0.143)	(0.745)	(1.820)
Observations	88	89	89	89	89	89
R-squared	0.352	0.290	0.364	0.654	0.693	0.304

Panel B: Press conference window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\lvert\Delta$ Similarity$\rvert$	0.005	0.035***	0.053*	0.270*	4.411*	5.021***
	(0.016)	(0.011)	(0.029)	(0.144)	(2.241)	(1.717)
Observations	89	89	89	89	89	89
R-squared	0.180	0.509	0.386	0.365	0.626	0.562

Note: The dependent variable is the realised variance of German sovereign yields at different maturities in the 30-minute window surrounding a monetary policy announcement. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.3: Changes in similarity and Italian sovereign yield variance

Panel A: Press release window

	(1) 1-year	(2) 2-year	(3) 5-year	(4) 10-year	(5) 20-year	(6) 30-year
$\lvert\Delta$ Similarity$\rvert$	-0.050	-0.016	-0.109	-0.740	-0.856	-2.112
	(0.050)	(0.043)	(0.114)	(0.647)	(0.529)	(1.550)
Observations	88	89	89	89	82	89
R-squared	0.459	0.552	0.210	0.183	0.334	0.274

Panel B: Press conference window

	(1) 1-year	(2) 2-year	(3) 5-year	(4) 10-year	(5) 20-year	(6) 30-year
$\lvert\Delta$ Similarity$\rvert$	-0.374	0.038	0.135**	1.255***	2.702***	4.915***
	(0.260)	(0.032)	(0.064)	(0.430)	(0.787)	(1.363)
Observations	88	89	89	89	82	89
R-squared	0.407	0.509	0.529	0.624	0.631	0.680

Note: The dependent variable is the realised variance of Italian sovereign yields at different maturities in the 30-minute window surrounding a monetary policy announcement. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.4: Changes in similarity and Spanish sovereign yield variance

Panel A: Press release window

	(1) 2-year	(2) 5-year	(3) 10-year	(4) 20-year	(5) 30-year
$\lvert\Delta$ Similarity$\rvert$	-0.021	-0.121	-0.339	-1.064	-1.040
	(0.028)	(0.210)	(0.350)	(0.974)	(0.974)
Observations	89	89	89	89	89
R-squared	0.274	0.344	0.326	0.412	0.237

Panel B: Press conference window

	(1) 2-year	(2) 5-year	(3) 10-year	(4) 20-year	(5) 30-year
$\lvert\Delta$ Similarity$\rvert$	-0.006	0.128***	0.155**	1.069**	2.764***
	(0.027)	(0.042)	(0.077)	(0.418)	(1.042)
Observations	89	89	89	89	89
R-squared	0.373	0.596	0.338	0.406	0.608

Note: The dependent variable is the realised variance of Spanish sovereign yields at different maturities in the 30-minute window surrounding a monetary policy announcement. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.5: Changes in similarity and French government bond absolute returns

Panel A: Press release window

	(1) 1-year	(2) 2-year	(3) 5-year	(4) 10-year	(5) 20-year	(6) 30-year
$\lvert\Delta$ Similarity$\rvert$	-0.049	-0.183	-0.157	0.257	0.172	-1.208
	(0.053)	(0.145)	(0.397)	(0.594)	(0.743)	(1.052)
Observations	88	89	89	89	89	89
R-squared	0.298	0.289	0.331	0.511	0.569	0.646

Panel B: Press conference window

	(1) 1-year	(2) 2-year	(3) 5-year	(4) 10-year	(5) 20-year	(6) 30-year
$\lvert\Delta$ Similarity$\rvert$	0.003	0.046*	0.130	0.376**	0.786**	1.607***
	(0.018)	(0.025)	(0.085)	(0.155)	(0.313)	(0.603)
Observations	67	89	89	87	89	89
R-squared	0.422	0.233	0.224	0.378	0.421	0.440

Note: The dependent variable is the absolute value of French sovereign bond returns at different maturities using high-frequency one-minute data. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.6: Changes in similarity and German bond absolute returns

Panel A: Press release window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\|\Delta$ Similarity$\|$	-0.079	-0.083	-0.576	-0.282	-0.661	-0.691
	(0.058)	(0.148)	(0.453)	(0.476)	(0.634)	(0.648)
	88	89	89	89	89	89
R-squared	0.193	0.361	0.323	0.524	0.543	0.087

Panel B: Press conference window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\|\Delta$ Similarity$\|$	0.001	0.027	0.053*	0.120	1.139**	2.694**
	(0.015)	(0.023)	(0.029)	(0.193)	(0.511)	(1.119)
Observations	79	87	89	89	89	88
R-squared	0.269	0.076	0.386	0.212	0.389	0.421

Note: The dependent variable is the absolute value of German sovereign bond returns at different maturities using high-frequency one-minute data. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $|\Delta$ Similarity$|$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.7: Changes in similarity and Italian bond absolute returns

Panel A: Press release window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\lvert\Delta$ Similarity$\rvert$	-0.007	-0.373	-1.509**	-2.832*	-3.145	-3.801
	(0.048)	(0.292)	(0.706)	(1.668)	(2.312)	(3.069)
Observations	88	89	89	89	82	89
R-squared	0.478	0.204	0.197	0.216	0.257	0.321

Panel B: Press conference window

	(1)	(2)	(3)	(4)	(5)	(6)
	1-year	2-year	5-year	10-year	20-year	30-year
$\lvert\Delta$ Similarity$\rvert$	-0.049	0.136	0.693*	1.390**	2.068***	2.344***
	(0.039)	(0.089)	(0.360)	(0.591)	(0.686)	(0.785)
Observations	73	89	89	89	82	88
R-squared	0.312	0.370	0.381	0.353	0.499	0.476

Note: The dependent variable is the absolute value of Italian sovereign bond returns at different maturities using high-frequency one-minute data. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.8: Changes in similarity and Spanish bond absolute returns

Panel A: Press release window

	(1) 2-year	(2) 5-year	(3) 10-year	(4) 20-year	(5) 30-year
$\lvert\Delta$ Similarity$\rvert$	0.147	-0.305	-0.188	-0.006	-4.532
	(0.122)	(0.243)	(0.630)	(0.795)	(2.833)
Observations	89	89	89	89	89
R-squared	0.221	0.309	0.571	0.509	0.330

Panel B: Press conference window

	(1) 2-year	(2) 5-year	(3) 10-year	(4) 20-year	(5) 30-year
$\lvert\Delta$ Similarity$\rvert$	0.051	0.188	0.525**	2.046***	2.329***
	(0.034)	(0.147)	(0.251)	(0.664)	(0.786)
Observations	89	87	88	89	87
R-squared	0.402	0.385	0.406	0.502	0.526

Note: The dependent variable is the absolute value of Spanish sovereign bond returns at different maturities using high-frequency one-minute data. Panel A presents the results obtained focusing on the press release window [13:30–14:00], while the press conference estimates [14:15–15:45] are presented in Panel B. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included in all specifications. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table D.9: Changes in similarity and Euro area stock market indices variance - alternative event window

Panel A: Press release window

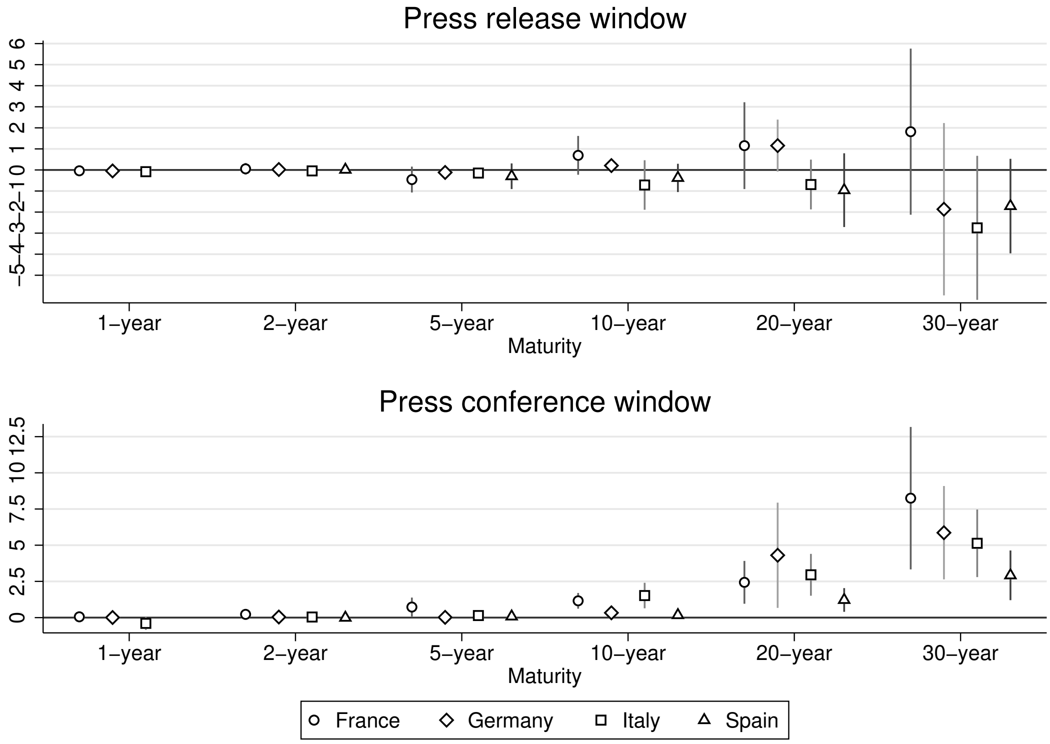
	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	0.745	-0.099	3.828	2.309	1.693	2.390
	(3.978)	(3.137)	(3.323)	(4.758)	(3.830)	(6.060)
Observations	89	89	89	89	89	89
R-squared	0.274	0.265	0.411	0.232	0.317	0.393

Panel B: Press conference window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	6.304***	5.619***	7.717***	7.210***	7.395***	12.845***
	(1.788)	(1.511)	(2.341)	(2.332)	(2.025)	(4.056)
Observations	89	89	89	89	89	89
R-squared	0.468	0.481	0.491	0.387	0.487	0.402

Note: The dependent variable is the realized variance of the stock market indices of major Euro area countries: CAC 40 for France, DAX for Germany, FTSE MIB for Italy, and IBEX for Spain, as well as the EURO STOXX50 (STOXX50E) and EURO STOXX Banks (SX7E) indices for European blue chip companies and banks in the 45-minute window around each event. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour prior to the release of a monetary policy decision. Year fixed-effects dummies are included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Figure D.1: Changes in similarity and European sovereign yields volatility during ECB announcements - alternative event window



Note: The figure show the coefficient of |Δ Similarity| in Eq. (3). The dependent variable is the realized variance of European sovereign yields around the ECB press release window [13:30–14:15] and ECB press conference window [14:15–16:00], respectively. Year fixed-effects are included. 90% confidence intervals are presented.

Table D.10: Changes in similarity on Euro area stock market indices returns - alternative event window

Panel A: Press release window

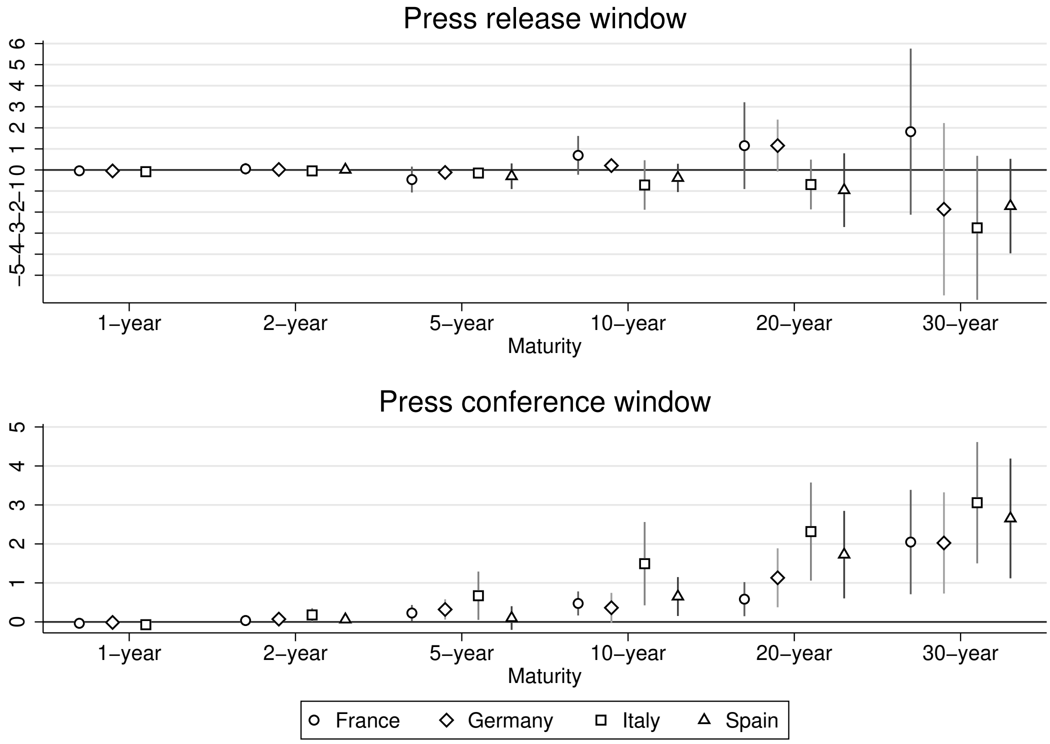
	(1)	(2)	(3)	(4)	(5)	(6)
	CAC 40	DAX	FTSE MIB	IBEX	STOXX50E	SX7E
|Δ Similarity|	2.337	1.087	3.474	5.047	3.332	8.438*
	(2.892)	(2.113)	(2.883)	(3.124)	(3.019)	(4.256)
Observations	89	89	89	89	89	89
R-squared	0.321	0.329	0.505	0.455	0.386	0.615

Panel B: Press conference window

	(1)	(2)	(3)	(4)	(5)	(6)
	CAC 40	DAX	FTSE MIB	IBEX	STOXX50E	SX7E
|Δ Similarity|	4.491**	4.984***	4.481**	2.414	3.727**	2.084
	(2.019)	(1.599)	(2.175)	(1.892)	(1.541)	(2.575)
Observations	89	89	89	89	89	89
R-squared	0.282	0.312	0.238	0.240	0.259	0.301

Note: The dependent variable is the absolute value of returns of the stock market indices of major Euro area countries, i.e. CAC 40 for France, DAX for Germany, FTSE MIB for Italy and IBEX for Spain, as well as the EURO STOXX50 and EURO STOXX Banks indices for the Euro area using high-frequency one-minute data. Returns are computed as the percentage variation in the median price between the 30-45 minutes following a press release and the 10-20 minutes before it. |Δ Similarity| is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Figure D.2: Changes in similarity and European sovereign yield returns during ECB announcements - alternative event window



Note: The figure show the coefficient of |Δ Similarity| in Eq. (3). The dependent variable is the absolute return of European sovereign yield around the ECB press release and press conference windows, respectively. Following (Altavilla et al., 2019), sovereign yield returns are computed as the percentage variation in the median price between the 30-45 minutes following an announcement and the 10-20 minutes before it. Year fixed-effects are included. 90% confidence intervals are presented.

Table D.11: Changes in similarity and Euro area stock market indices variance - VSTOXX robustness

Panel A: Press release window

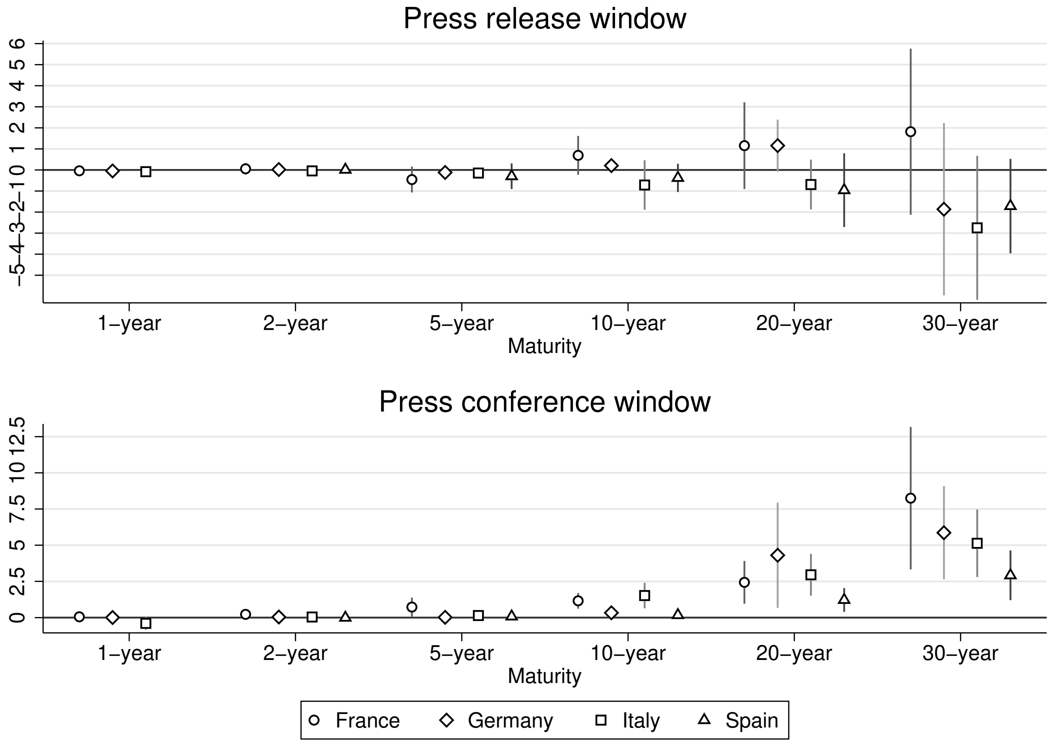
	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\|\Delta$ Similarity$\|$	0.977	0.294	3.700	2.733	2.016	3.628
	(3.720)	(2.929)	(3.132)	(4.458)	(3.613)	(5.312)
VSTOXX	0.035	0.028	0.014	0.020	0.035	0.032
	(0.027)	(0.018)	(0.019)	(0.023)	(0.027)	(0.045)
Observations	89	89	89	89	89	89
R-squared	0.306	0.298	0.387	0.242	0.350	0.415

Panel B: Press conference window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\|\Delta$ Similarity$\|$	5.010***	4.303***	6.284***	5.813***	5.850***	10.446***
	(1.435)	(1.212)	(2.080)	(1.979)	(1.613)	(3.472)
VSTOXX	0.109***	0.093***	0.105***	0.107***	0.125***	0.193***
	(0.023)	(0.018)	(0.023)	(0.025)	(0.024)	(0.049)
Observations	89	89	89	89	89	89
R-squared	0.601	0.599	0.607	0.489	0.620	0.519

Note: The dependent variable is the realized variance of the stock market indices of major Euro area countries: CAC 40 for France, DAX for Germany, FTSE MIB for Italy, and IBEX for Spain, as well as the EURO STOXX50 (STOXX50E) and EURO STOXX Banks (SX7E) indices for European blue chip companies and banks in the 30-minute window around each event. $|\Delta$ Similarity$|$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour prior to the release of a monetary policy decision. Year fixed-effects dummies are included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Figure D.3: Changes in similarity and European sovereign yields volatility during ECB announcements - VSTOXX robustness



Note: The figure show the coefficient of |Δ Similarity| in Eq. (3). The dependent variable is the realized variance of European sovereign yields around the ECB press release window [13:30–14:00] and ECB press conference window [14:15–15:45], respectively. Year fixed-effects are included. 90% confidence intervals are presented.

Table D.12: Changes in similarity on Euro area stock market indices returns - VSTOXX robustness

Panel A: Press release window

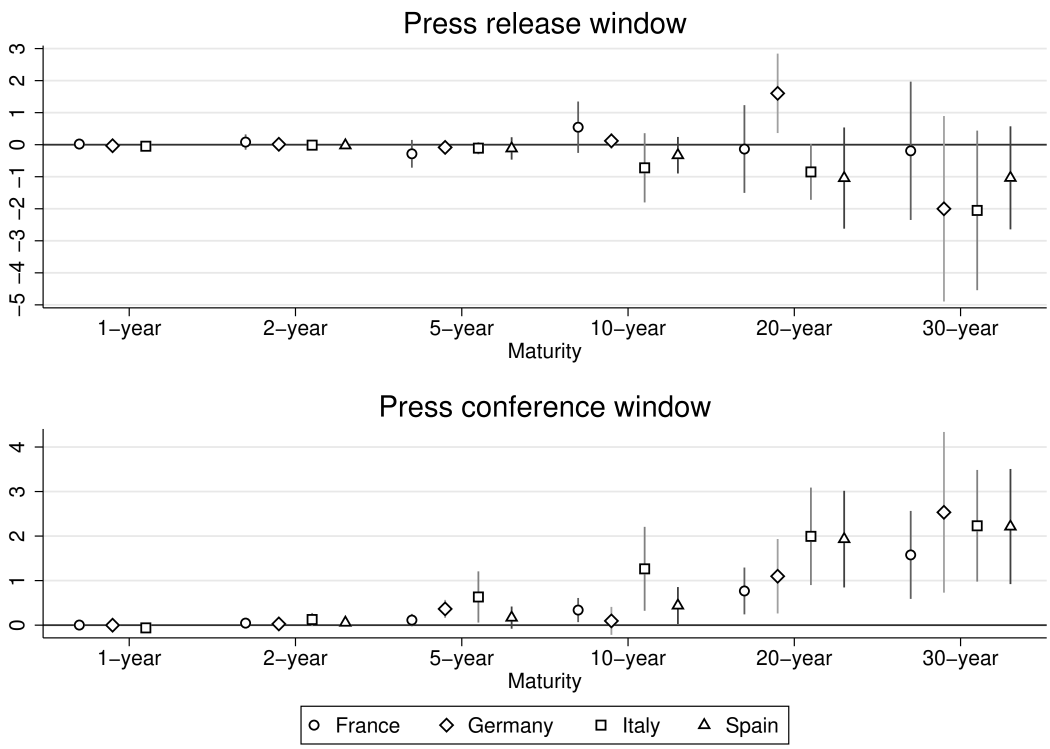
	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	3.525	1.798	6.050	5.800*	4.302	10.638*
	(2.891)	(2.193)	(4.117)	(3.340)	(3.175)	(5.625)
VSTOXX	0.043	0.032*	0.045	0.031	0.042	0.029
	(0.026)	(0.017)	(0.046)	(0.028)	(0.029)	(0.061)
Observations	89	89	89	89	89	89
R-squared	0.412	0.411	0.492	0.470	0.439	0.579

Panel B: Press conference window

	(1) CAC 40	(2) DAX	(3) FTSE MIB	(4) IBEX	(5) STOXX50E	(6) SX7E
$\lvert\Delta$ Similarity$\rvert$	4.002**	4.679***	3.841**	2.087	2.230**	1.689
	(1.580)	(1.472)	(1.581)	(1.427)	(1.059)	(1.902)
VSTOXX	0.050**	0.041**	0.034*	0.033*	0.046***	0.047*
	(0.020)	(0.019)	(0.019)	(0.017)	(0.017)	(0.026)
Observations	89	89	89	89	89	89
R-squared	0.282	0.312	0.238	0.240	0.259	0.301

Note: The dependent variable is the absolute value of returns of the stock market indices of major Euro area countries, i.e. CAC 40 for France, DAX for Germany, FTSE MIB for Italy and IBEX for Spain, as well as the EURO STOXX50 and EURO STOXX Banks indices for the Euro area using high-frequency one-minute data. Returns are computed as the percentage variation in the median price between the 15-25 minutes following a press release and the 10-20 minutes before it. $\lvert\Delta$ Similarity$\rvert$ is the absolute change in the similarity index between Twitter traffic and the policy announcement in the hour post as compared to the hour before the release of a policy decision. Year fixed-effects dummies are included, but not reported. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Figure D.4: Changes in similarity and European sovereign yield returns during ECB announcements - VSTOXX robustness



Note: The figure show the coefficient of |Δ Similarity| in Eq. (3). The dependent variable is the absolute return of European sovereign yield around the ECB press release and press conference windows, respectively. Following (Altavilla et al., 2019), sovereign yield returns are computed as the percentage variation in the median price between the 15-25 minutes following an announcement and the 10-20 minutes before it. Year fixed-effects are included. 90% confidence intervals are presented.