HONG KONG INSTITUTE FOR MONETARY AND FINANCIAL RESEARCH

UNCERTAIN TALKING AT CENTRAL BANK'S PRESS CONFERENCE: NEWS OR NOISE?

Nan Hu and Zexi Sun

HKIMR Working Paper No.04/2021

February 2021

HONG KONG ACADEMY OF FINANCE
金融學院

Hong Kong Institute for Monetary and Financial Research
香港貨幣及金融研究中心
(a company incorporated with limited liability)

Uncertain Talking at Central Bank's Press Conference: News or Noise?

Nan Hu
Hong Kong Institute for Monetary and Financial Research

Zexi Sun
Goethe University Frankfurt

February 2021

Abstract

We show that the uncertain talking at the European Central Bank's press conference raises stock returns of the Eurozone. By decomposing the uncertain talking into two types, we find that the positive returns are driven by one type, the average meaning of which is represented by words "risk", "uncertainty", "volatility", and "turbulence". This type is related to uncertain and unfa-vorable environments, which is negative news that lowers returns. However, an increase in this type also signals expansionary monetary policies in the future, which is positive news that raises returns. We find further evidence to imply that the latter channel dominates.

* Email: Hu: nhu@hkma.gov.hk and Sun: zsun@wiwi.uni-frankfurt.de

* We greatly appreciate the comments and suggestions from Michael Binder, Varadarajan Chari, Michael Ehrmann, Jan Pieter Krahnen, Michael McMahon, Paul McNelis, Emmanuel Moench, Stefano Nardelli, John Rogers, Maik Schmel-ing, Tatevik Sekphosyan, Matthew Shapiro, Peter Tilmann, Christian Wagner, seminar participants at Goethe University Frankfurt, International Monetary Fund, European Central Bank, as well as conference participants at the European Commission-JRC Big Data and Economic Forecasting Workshop, Georgetown Center for Economic Research Biennial Conference 2019, the 21st INFER Annual Conference, the 10th International Workshop on Theoretical and Experimental Macroeconomics, the 34th Annual Congress of the European Economic Association, and the 50th Anniversary Confer-ence of the Money, Macro and Finance Group. The views expressed in this paper are those of the authors, and do not necessarily reflect those of the Hong Kong Monetary Authority, Hong Kong Academy of Finance, Hong Kong Institute for Monetary and Financial Research, its Council of Advisers, or the Board of Directors. All remaining errors are our own.

Disclosure Statement

Nan Hu

I am an economist employed by the Hong Kong Institute for Monetary and Financial Research (HKIMR), an affiliation of the Hong Kong Monetary Authority. The paper is a working paper of the HKIMR. When I wrote the paper, I was employed as a project officer by the International Monetary Fund (2018, 2019), and a research assistant by the SAFE research center of the Goethe University Frankfurt (2018). The International Monetary Fund and Goethe University Frankfurt do not have an interest in the research in this paper.

An earlier version of the paper was a part my master thesis at the Goethe University Frankfurt. My master and doctoral supervisor, Professor Maik Schmeling, reviewed the paper. My present employee HKIMR has the rights to review the paper.

The student traveling funding of Goethe University Frankfurt's has funded our participation in the 34th Annual Congress of the European Economic Association, and the 50th Anniversary Conference of the Money, Macro and Finance Group. The Bank of Canada has provided full financial support for my trip to the 10th International Workshop on Theoretical and Experimental Macroeconomics. The SAFE data center of Goethe University Frankfurt provides access to the daily data used in the paper. The Euro Area Monetary Policy Event-Study Database offers free intraday data for this research.

I have no other potential conflicts to disclose.

Zexi Sun

I was employed as research assistant by the Goethe University Frankfurt when I wrote this paper. The Goethe University Frankfurt does not have an interest in the research in this paper.

My doctoral supervisors, Professor Michael Binder and Professor Volker Wieland, reviewed this paper prior to submission.

I have no other potential conflicts to disclose.

1 Introduction

In an era when conventional monetary policy tools are less effective, investors become more attentive than ever to the textual and verbal information released by the central bank. Studies show that the linguistic features of central bank announcements, including the tightening/easing inclination (Ehrmann and Fratzscher, 2007), the optimistic/pessimistic attitude (Born et al., 2014), and the positive/negative tone (Schmeling and Wagner, 2019) all significantly affect the asset prices.

Central bankers also turn increasingly cautious when framing their statements. Though the target of communication is to decrease uncertainty around banks' goals and strategies (Bernanke, 2007), the central banks are faced with uncertainty themselves. Unavoidably, central banks use uncertain expressions in their communications, but the effect of such uncertain talking is unknown. In the finance literature, imprecise information released by private firms is regarded as noises and is negatively correlated with stock market returns (Li, 2005; Epstein and Schneider, 2008; Ai, 2010; Brevik and d'Addona, 2010). While the roles played by central banks and the firms in the market are very different, empirical evidence on central banks is surprisingly scarce.

In this paper, we quantify the uncertain talking at the European Central Bank (ECB)'s press conference and examine the effect of the uncertain talking on the financial market. We find that the uncertain talking in presidents' answers constitutes positive news that raises stock returns. A one percent increase in the amount of unexpected uncertain talking raises the return of Euro STOXX 50 (SX5E) index by around 20 basis points. The positive effect is robust to the inclusion of control variables and alternative choices of uncertain talking measures. It remains stable in both intraday and daily time windows around the press conference.

We use the well-recognized uncertain word dictionary of Loughran and McDonald (2011) to build up the uncertain talking measures. However, the meanings and usages of the words from this dictionary are not identical. Consequently, there can be different types of uncertain talking. To rationalize the above findings, we adopt a novel approach that combines natural language processing (NLP) and machine learning (ML) techniques to classify uncertain words into two groups. The classification is based on the contexts of the ECB documents and thus accurately reflect how these words are used in the ECB statements. The first group contains words such as *risk*, *uncertainty*, *volatility*, and *turbulence*, which are related to uncertain and unfavourable economic and financial environments faced

by the ECB. The second group contains words such as *perhaps*, *almost*, *probably*, and *assumption*, which are related to uncertain attitude of the ECB president. Accordingly, we name the decomposed uncertain talking measures constructed from these two groups as the *environmental* uncertain talking and *attitudinal* uncertain talking measures, respectively. We find that positive returns are mainly driven by the environmental uncertain talking.

To explain the positive reaction of stock returns to the environmental uncertain talking, we propose a two-channel interpretation and apply a conceptual framework as in Bernanke and Kuttner (2005) to analyze its influence. The environmental channel works when investors update their expectations on the economic and financial conditions. The policy channel works when investors update their expectations of the future monetary policy stance. An increase in the environmental uncertain talking suggests a more uncertain and unfavorable environment. Through the environmental channel, it informs the investors of worsened economic and financial environments, which is negative news that reduces stock returns. Through the policy channel, it signals expansionary monetary policy moves in times of uncertainty and distress, which is positive news that raises stock returns. The overall positive sign implies that the policy channel dominates.

Why does the policy channel prevail? An explanation might be that the central bank announcement is one of the many sources from which market participants obtain new information about economic and financial conditions. However, it is the only source upon which they can revise their beliefs on *central bankers' views* of these conditions. Thus, a change in the environmental uncertain talking is more likely to revise investors' expectations of monetary policy paths through the policy channel, rather than their expectations on the conditions themselves through the environmental channel. A weak environmental channel is also supported by Bauer and Swanson (2020), in which the authors provide direct evidence to show that professional forecasters rarely revise their forecasts on fundamentals in response to the announcements from the Federal Reserve.

We provide empirical evidence on the existence of the policy channel. We show that both the EONIA-based overnight indexed swap (OIS) rates and the nominal German government bond yields fall in response to an increase in the environmental uncertain talking. Such an effect is significant for the OIS rates and bond yields with one year up to 20 years of maturities. This implies that the environmental uncertain talking changes market-based expectations of the medium- to long-term monetary policy paths.

We further support the two-channel interpretation by analyzing the effectiveness of the environmental and future policy channels under various sub-samples. The environmental uncertain talking is a verbal signal (the "soft information") that influences investors' expectations through two channels. The ECB sometimes also releases numeric signals (the "hard information") such as monetary policy decisions and official projections. They can also affect investors' expectations of the environment and monetary policy stance. According to the Bayes' rule, agents will place more weight on signals with higher precision when revising their expectations. Hautsch and Hess (2007) also find that asset prices respond significantly stronger to precise news than imprecise news. Arguably, the numerical signals such as policy decisions or official projections are more precise than verbal information. Therefore, we expect that the effectiveness of the two channels are affected by the release of numeric policy decisions and official projections.

We find that, when the ECB announces monetary policy changes on press conference days, the impact of the environmental uncertain talking turns out to be insignificant. Conversely, when ECB provides macroeconomic projections, the impact of the environmental uncertain talking is strengthened. This confirms our prediction: When a policy action is released, the policy channel of the environmental uncertain talking is weakened and no longer influences the market significantly. Likewise, official projections reduce the effectiveness of the environmental channel, hence enhance the relative strength of the policy channel, and the positive reaction of stock returns becomes even stronger. The results also suggest that the soft and hard information in our context are substitutes.

Our paper contributes to the current literature by discovering a significant positive impact of central bankers' uncertain talking on stock returns and explaining this effect through the dominance of the policy channel over the environmental channel. These findings have important implications for the communication policies of the central banks and investment strategies of the investors.

Our findings point out an overlooked issue in theoretical models on imprecise information. While some asset-pricing models suggest that uncertain signals released by private firms have a negative impact on stock returns, our results show that the uncertain signals released by central banks lead to higher stock returns. The crucial difference is that, from the uncertain talking of the policymaker, market participants infer not merely the future states, but also state-dependent future policy paths. Hence, it is essential to distinguish between uncertain public signals released by private firms and by the policymaker.

Finally, we contribute methodologically by proposing a new approach to accurately and objectively decompose a pre-defined word list into different dimensions, according to the context in which the words are actually used. This method can be applied to textual analyses for research on other topics.

The remainder of this paper is organized as follows. Section 2 introduces the literature. Section 3 measures the uncertain talking at the ECB's press conferences. Section 4 discusses the textual and numeric data used in the empirical analysis. Section 5 addresses the estimation strategy. Section 6 presents the regression results, including the robustness test. Section 7 identifies two types of uncertain talking. Section 8 explains the positive reaction of stock returns through the two-channel interpretation. Section 9 provides evidence to support the interpretation. Section 10 concludes.

2 Related Literature

Our study is related to the asset pricing literature on information quality. Most studies in this branch find that imprecise information increases risk premium and consequently lowers the equity price, which includes theoretical studies such as Epstein and Schneider (2008) and empirical research such as Loughran and McDonald (2011), Loughran and McDonald (2013), and Dzieliński et al. (2017). Unlike existing studies, our paper shows that imprecise information released by policymakers affects the financial markets differently from the imprecise information released by firms. It is because market participants are attentive to not only the facts, but also how the policymakers read these facts, as the latter signals central bank's monetary policy stance in the future.

Our paper is connected with researches exploring how the quantified textual information interacts with the financial markets. Tetlock (2014) and Loughran and McDonald (2016) provide comprehensive surveys on this subject. Recent studies apply textual analysis to a broader range of topics such as hedging risks contained in news (Engle et al., 2020), measuring innovation and product differentiation (Kostovetsky and Warner, 2020), and exploring the similarity of the language used in firms' financial reports as well as its implication on firms' operations (Cohen et al., 2020). We start by employing the word list from Loughran and McDonald (2011) to obtain the uncertain talking measure. Then we propose a novel method to detect different types of uncertain talking by using the contextual information of the ECB documents.

We explain the significant response of stock returns through the environmental channel and the pol-

icy channel. These two channels have been extensively discussed in both theoretical and empirical works (e.g. Nakamura and Steinsson, 2018; Andrade et al., 2019). While some of these works argue that policy decisions can change people's expectations of the fundamentals, our results indicate that policymakers' views on economic conditions can change investors' expectations of the future policy stance.

Finally, our paper belongs to the literature investigating the impact of central bank's policies on stock markets. Kuttner (2001) and Bernanke and Kuttner (2005) pioneered in constructing the monetary policy surprises and investigating their effects on US equity prices. Lucca and Moench (2015) and Cieslak et al. (2019) find connections between policy moves and excess returns in the stock market. Schmeling and Wagner (2019) explores how the central bank's negative tone moves asset prices. We contribute to this branch by revealing a significant and positive effect of uncertain talking on stock returns and by providing intuitive explanations and supporting evidence to explain it.

3 Measuring the uncertain talking

3.1 The Structure of the ECB's Press Conference

This section discusses how to measure the uncertain talking at the ECB's press conference. Although any verbal or textual messages released by the ECB can be considered part of its communication policies, we only focus on the impact of the ECB presidents' talks at the press conference for four reasons. First, the ECB's press conference is held within an hour after it releases the monetary policy decisions, at which a detailed explanation of the rationale behind the decisions is provided. Thus, this is the most direct and fundamental source for understanding the ECB's policy considerations and learning about its judgment on the economic and financial conditions. Second, this is the only event for which the ECB offers regular live broadcasts on its official YouTube channel. Besides, both Bloomberg and Reuters, two of the world's largest information providers, offer comprehensive live coverage of the ECB's press conferences on their platforms. It is thus reasonable to assume that, among all the events held by the ECB, the press conference is the one that attracts the most attention and can lead to the most significant reaction in the financial market. Third, unlike other communication channels such as speeches and interviews, market participants already have formalized clear and consistent expectations about the date, speaker, and structure of the press conference. Thus, if

any, asset price movements around the press conference will mostly reflect unexpected changes in the content of the statement. Finally, the press conference of the ECB has been held regularly since its foundation, and thus it provides us with an ample number of observations to gauge its impact empirically.

The press conference always takes place on a trading day, mostly on Thursdays, and sometimes on Wednesdays, which is ideal for us to inspect the financial market response within the narrow event time window around the press conference. It starts at 14:30, 45 minutes after the press release of policy decisions, and ends at approximately 15:30. It is usually split into two sections: For the first 10-15 minutes, the president reads out a prepared Introductory Statement to clarify policy decisions. Then, for the remaining 40-45 minutes, the president (and occasionally the vice presidents) answers questions from journalists in the Question and Answer (Q & A) session. Figure 1 shows the timeline of a typical press conference day. For each press conference transcript, we independently measure

Figure 1
Timeline of Press Conference Day

This Figure shows the timeline of a typical press conference day. The three time windows are defined in Altavilla et al. (2019).

and analyze the effect of the uncertain talking in (1) the Introductory Statement, (2) the questions from the journalists, and (3) the answers from the president. The reason for separating the uncertain talking in the questions from other parts is obvious, as the former reflects the reporters' attitude but not policymakers. The reason for distinguishing between the uncertain talking in the Introductory Statement and answers is worthy of a further discussion. As the Introductory Statement is drafted in advance, its content is strictly controlled by the ECB staff. Therefore, it can only convey the *intended amount of uncertain talking* that the policymakers want to show to the public. On the contrary, it is more difficult for the ECB president to control the language used in answers since the questions that journalists raise are not communicated with the ECB staff before the press conference. Therefore, the

uncertain talking in answers more reflects the *perceived amount of uncertain talking* the policymakers have in mind. Thus, the uncertain talking in these two parts can result in very different interpretations in the market, which are proved by the empirical results in the next sections.

3.2 Defining the Uncertain Talking Measure

To quantify the uncertain talking, we apply an automated dictionary method to count the appearance of words from a pre-defined uncertain word list (that is a "*dictionary*"). This dictionary method is the standard approach for calculating a particular metric (which in our case is the measure of uncertain talking) from unstructured textual data and is adopted in other studies to measure the tone from central bank's policy statements (e.g. Hansen and McMahon, 2016; Schmeling and Wagner, 2019)

The dictionary we employ is the uncertain word list built by Loughran and McDonald (2011) to construct the uncertain talking measure. This dictionary has been used to measure the extent of uncertainty in the written (Loughran and McDonald, 2013) and verbal (Dzieliński et al., 2017) information from private firms. Recent studies have also applied this dictionary to analyze the impact of uncertainty in central bank statements (Jegadeesh and Wu, 2017; Picault and Renault, 2017).

For a press conference at date t, the uncertain talking measure in part $K \in \{N, Q, A\}$ is defined as

$$UncK\%_t = \frac{\#(Uncertain\ words\ in\ part\ K)_t}{\#(All\ words\ in\ press\ conference\ transcript)_t} \times 100, \tag{1}$$

where N, Q, and A stand for the Introductory Statements, questions from journalists, and answers from the ECB president, respectively. The numerator is the number of words from the uncertain word dictionary in part K, and the denominator is the total number of words in the press conference transcript.

In the equation above, we take a relative measure (i.e., the percentage of words) rather than an absolute measure (i.e., the number of words), so as to isolate the impact of uncertain talking from the effect of text length, which itself is usually treated as a readability measure and will be controlled for. In the robustness check, we show that taking an absolute measure does not alter our conclusion.

Besides, the same scaling parameter (i.e., $\#All\ words\ in\ press\ conference\ transcript$) is used for calculating uncertain talking measures in three parts. This guarantees that the difference in the impacts

of uncertain talking of different parts is not a result of the variation of this parameter. We also show in the robustness check that the main result remains consistent when we use a unique scaling parameter for each part[1].

The dictionary method has several unique advantages over the conventional approach of manually scoring the text: It is more objective and can be easily replicated, extended, and validated with minimal cost. However, it also has two limitations: First, since the dictionary is pre-compiled, it ignores the contextual information of the text. Second, the precise meanings and usages of words in the dictionary can be unalike.

We present the three examples of sentences with uncertain words in Table I. Although these words are all related to specific dimensions of uncertainty, they are used in different situations, and their meanings are not similar. In the first example, the modal word *maybe* is used to indicate the policymaker's uncertain view of a policy strategy. In contrast, the nouns *turbulence* and *volatility* in the second example do not seem to convey an uncertain attitude of the speaker. Instead, they are purely descriptions of the uncertain circumstances the ECB faced.

Taking both the advantages and limitations of the dictionary method into account, to begin with, we study the effect of uncertain talking with a measure that is built upon the standard uncertain word dictionary from Loughran and McDonald (2011). In the second half of the paper, we decompose the overall measure into different dimensions, based on how the uncertain words are actually used. Then we examine which dimension propels significant responses in asset prices.

4 Data

4.1 Textual Data

To obtain the uncertain talking measure, we download the ECB's press conference transcripts from its website[2]. In total, we collect 194 ECB press conference transcripts from Dec 2000 to Sep 2018[3].

[1] Another issue worth noting is that we assign each word the same weight when building up this measure. An alternative choice is the TF-IDF weighting scheme, which assigns words appearing less often with higher weights, in order to "highlight words that are distinctive of the current text" Goldberg (2017). However, as emphasized by Schmeling and Wagner (2019), this would imply a hindsight bias as this scheme requires the data to be used twice. Even so, we have checked that choosing the TF-IDF weighting scheme does not alter our conclusion.

[2] https://www.ecb.europa.eu/press/pressconf

[3] The ECB started the press conference from June 1998. However, we only begin our analysis of the press conference transcripts from December 2000, because the ECB went through significant revisions of its monetary and communication

Table I
Sentences with Uncertain Words in the ECB's Press Conference Transcripts

This table provides three examples of sentences that contain uncertain words and are spoken by the ECB presidents at press conference to answer questions in the Q&A session. Source: ECB.

President	Date	Subject	Example
Duisenberg	4/3/2003	Policy strategy	*... maybe that will be final, maybe we will need another discussion two weeks later ...*
Trichet	8/7/2008	Financial market	*... it's an ongoing, important market correction with episodes of turbulence and high levels of volatility ...*
Draghi	7/6/2018	Unconventional Policy	*... you have the uncertainty component in the term structure; this of course tends to be more variable and shifts with risk perceptions ...*

During this period, the ECB's press conferences always took place after a press release of the monetary policy decision. However, there were a few times when only a policy decision was released, but not followed by a press conference[4].

Each transcript is separated into three parts: the Introductory Statement (N), questions (Q), and answers (A). We then employ a standard text-cleaning process on each part, which involves converting words to lower case, lemmatizing and stemming words, and removing numbers, web addresses, non-English words, and other stop words that do not convey essential information. For each Introductory Statement, we further remove the welcoming and concluding remarks that bear no economic meaning. In the appendix we provide all the cleaning details.

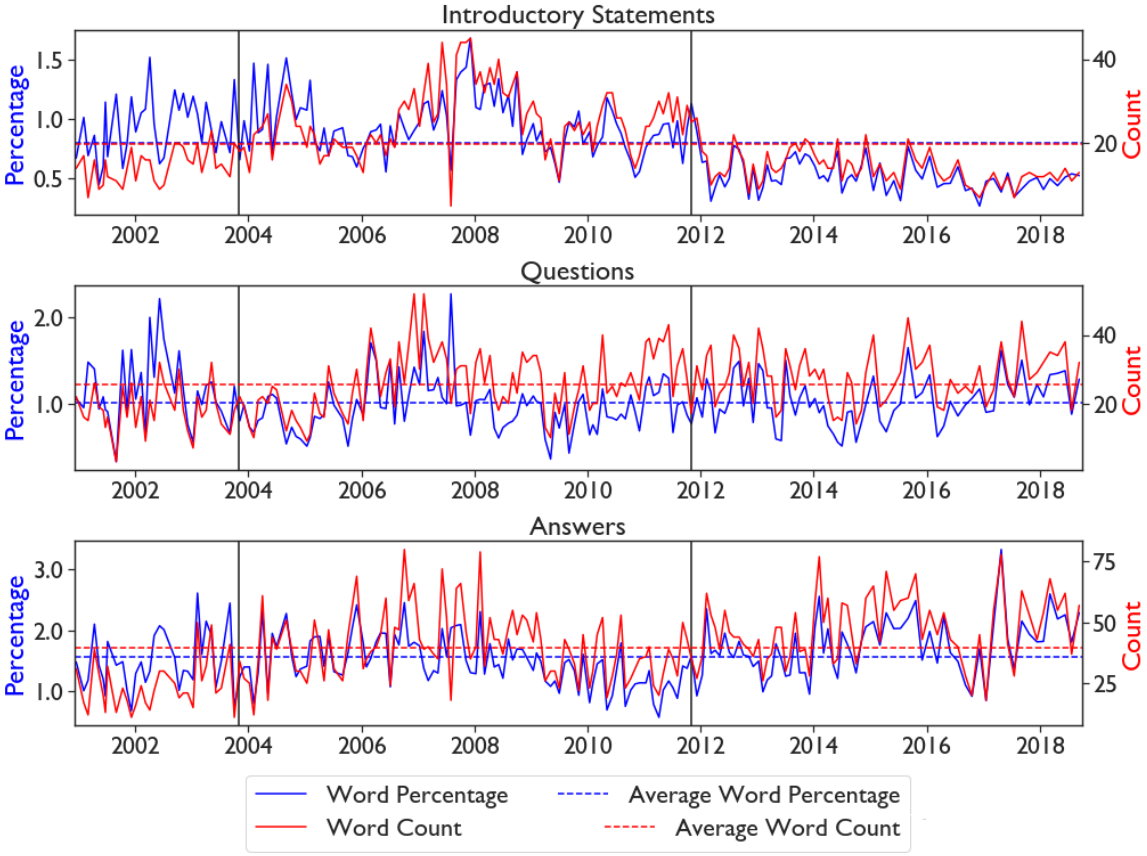
After finishing the text-cleaning procedure, we calculate the uncertain talking measures in part $k \in \{N, Q, A\}$ based on (1). The summary statistics of the uncertain talking measures are displayed in Table II. On average, there are approximately 19.69, 38.22, and 23.48 uncertain words in the Introductory Statement, questions, and answers, respectively. These uncertain words account for roughly 0.79%, 1.50%, and 0.93% of the total word count in a press conference transcript. The correlation between any two of the three uncertain talking measures is small. The negative correlation (-0.213) between the uncertain talking measures in the Introductory Statement and answers further proves the necessity of separating these two parts, while the low correlation (0.079) between the uncertain talking measures in the questions and answers provides little support to the conjecture that the president will be inclined to talk uncertainly when answering uncertain questions.

practices in 1998-2000. Also, we do not include the ECB's press conferences on topics irrelevant to its monetary policy.

[4]Before 2001, the policy decision was released every two weeks, but the press conference was held every four weeks, meaning that one press conference was held for every two policy decisions. Before 2006, during the summer break, the ECB only released policy decisions without holding the press conference. The last exception was October 8th, 2008, when the ECB took an emergency action with other major central banks to fight the global financial crisis.

Figure 2
Time-Series Plots of the Uncertain Talking Measures

This figure shows the time-series plots of the uncertain talking measures in three parts of ECB's press conference. The sample period is from Dec 2000 to Sep 2018. The red lines and blue lines connect the total number of uncertain words and the percentage of uncertain words (calculated based on (1)) in each transcript, respectively. The dashed lines show the average levels of the corresponding series in the solid lines. The black vertical lines signify the change in the ECB presidency (Wim Duisenberg, 1999-2003; Jean-Claude Trichet, 2003-2011; Mario Draghi: 2011-2019). Sources: ECB, Authors Calculations.



We plot the time-series of the uncertain words in each part of the transcript in Figure 2. In the first subplot, there is a notable downward trend in the count and percentage of uncertain words in the Introductory Statement after 2008, which becomes even more evident during Mario Draghi's presidency. This drop is unlikely to signal subdued economic and policy uncertainties in the Euro Area during this period. Instead, it suggests that the ECB has been controlling the uncertain language in its Introductory Statement from 2011-2018 on purpose, possibly attempting to anchor expectations during this turbulent period. This downward trend cannot be found in the count or percentage of uncertain words in the answers. Controlling the language used in answering questions without full preparation is more challenging than controlling the language used in a well-prepared statement.

We present an exhaustive list of the sentences containing uncertain words in the ECB's press conference transcripts online, which is described in Appendix.

Besides the uncertain talking measure, we also count the total number of words in each transcript and calculate the negativity measure in each part of the transcript based on the negative word dictionary from Loughran and McDonald (2011)[5]. Both the word count and negative measures will be treated as control variables in the regression.

4.2 Financial and Economic Data

Our baseline analysis concentrates on the reaction of the Euro STOXX 50 (SX5E) index. As one of the Eurozone's benchmark indices, SX5E represents the stock prices of fifty leading super-sector leaders in the region and serves as an underlying asset for many ETFs, futures, and options.

We are interested in the intraday returns of the SX5E index in the narrow time window around the ECB's press conference. This high-frequency data is from the Euro Area Monetary Policy Event-Study Database (EA-MPD) constructed by Altavilla et al. (2019). The EA-MPD contains the price changes in three time windows, including the press release window, press conference window, and monetary event window, on a press conference day. The price change of the SX5E index in the press conference window is the focus of our study[6]. Figure 1 displays the time windows on a press conference day.

We also investigate the impact of uncertain talking on the daily returns of a broad spectrum of Euro-zone stock market indices. This includes Europe STOXX Large 200, Europe STOXX Median 200, Europe STOXX Small 200, MSCI Europe, MSCI Europe Growth, and MSCI Europe Value indices, which are all obtained from Datastream. To test whether the uncertain talking has a particular influence on a specific country or industry, we further collect the daily returns of ten MSCI country indices and eighteen Euro STOXX total market industry indices. All the daily returns are defined as the close-to-close percentage change of index prices.

[Table II about here]

We further collect the high-frequency price changes of other assets, including the EONIA-based

[5]Notice that there is an overlap between the uncertain and negative words. Since these repeated words only take up a small fraction of the total number of the uncertain or negative words in the transcript, we remove these overlapping words from the negative word list to focus on the impact of uncertain talking. Keeping these words in the negative word list does not alter our conclusion.

[6]The authors do not construct the price change by selecting the price at a particular tick. Instead, they choose the median price in a short time interval as the starting and ending quotes in each window. As a result, the length of the press conference window varies between 75 to 95 minutes.

overnight indexed swaps and German government bonds, from the EA-MPD. We retrieve the changes in the ECB's three key policy rates, its forward guidance statements, and its plan for the asset purchasing program from its official website. Finally, we download the ECB staff projections, the Economic Policy Uncertainty index (Baker et al., 2016), and data for the US initial jobless claims from Datastream. Table II provides the summary statistics of the main variables used in our analysis.

5 Estimation strategy

We apply a standard event-study approach to examine whether there exists a causal relationship between the uncertain talking at the ECB's press conference and Eurozone stock market returns. The regression equation takes the form

$$R_t = \beta_0 + \beta_1 \Delta UncK\%_t + \boldsymbol{\gamma}' \boldsymbol{x}_t + \varepsilon_t, \tag{2}$$

In our baseline analysis, R_t refers to the percentage return of the SX5E index in the press conference window. $\Delta UncK\%_t$ is the change of the uncertain talking measures from the previous to the current press conference in part K, which can be the Introductory Statement ($K = N$), questions ($K = Q$) or answers ($K = A$).

We build our analysis based on the efficient market hypothesis, that is, asset prices should respond only to the unanticipated information. Then stock returns should react only to the surprised component, that is the second term in the equation below:

$$UncK\%_t = UncK\%_t^{expected} + UncK\%_t^{surprised}. \tag{3}$$

Similar to the set-up in Schmeling and Wagner (2019), in our baseline specification, we assume that the expected component is the first lag of the uncertain talking measure; that is, the investors' anticipated level of uncertain talking in the current press conference equals the level they observed in the last press conference. Thus, as shown in the equation below, the surprise component becomes the first difference of the uncertain talking measure:

$$\Delta UncK\%_t^{surprised} = UncK\%_t - UncK\%_t^{expected} \approx UncK\%_t - UncK\%_{t-1} = \Delta UncK\%_t. \tag{4}$$

We also exploit other possibilities. On the one hand, it might be that investors have no prior expectation of the level of uncertain talking so that the expected component is zero, and the surprise component is equal to the level of uncertain talking in the current press conference. On the other hand, investors can predict the current level of uncertain talking based upon a forecasting regression using past information, in which case the unanticipated component is the residual. In the robustness check, we examine whether the main findings remain consistent under these two alternative assumptions.

We include a set of control variables $\boldsymbol{x}_t$. Since the answers from ECB presidents are unlikely to be affected by the price movements during the press conference, the likelihood of reverse causality is relatively low[7]. Nevertheless, there might still be forces that affect both the amount of uncertain talking and the movement of stock prices. To isolate the influence of uncertain talk from other effects, we include the following control variables.

First, we control for several textual features of the press conference transcript. Schmeling and Wagner (2019) find that the positive/negative tone of the policymaker can significantly move the SX5E index. Coenen et al. (2017) indicate that the complexity of the ECB statement has a significant impact on the option-implied volatility of the SX5E index. To exclude these potential influences, we control for (1) the first difference of the negativity measures in three parts of the transcript ($\Delta NegK\%_t$ for $K \in \{N, Q, A\}$), as well as (2) the change in the total number of words in a transcript ($\Delta TWords$) as a proxy of the readability measure[8].

Second, to separate the effects of hard information (i.e., quantitative numeric signals such as policy action) from soft information (i.e., qualitative verbal signals including uncertain talking), we control for a set of policy decisions and announcements on press conference days. These controls include (1) changes in the marginal refinancing operation rate (ΔMRO_t), which represents the ECB's conventional policy decision[9]; (2) A dummy variable (FG) representing the ECB's unconventional policy

[7] In the press conference held by the Federal Reserve on July 31st, 2019, a journalist told Chairman Jay Powell that "As this press conference has gone underway, markets have declined." In this example, a reverse causality could exist if Powell's answer was based on the real-time development of the financial markets. However, we do not find any similar expressions in questions raised in the ECB's press conferences in our sample period.

[8] It should be noted that several studies (e.g. Tobback et al., 2017) show that the dovish/hawkish inclination is also an important dimension of central banks' statements. However, since there is no well-established dictionary for calculating this metric, we need to either build up a word list according to some subjective criteria, or to manually annotate the polarity of each sentence in the statement, or to train a machine learning model based on manually coded examples, which are all beyond the scope of this paper.

[9] The marginal refinancing operation rate is the middle of the three official interest rates. ECB regularly adjusted the MRO rate before it hit the zero lower bound in 2016. The implementation of the new official interest rate usually

moves, which equals one when there is a forward guidance statement in the policy statement and zero otherwise; (3) market-based measures of monetary policy surprise, measured as the first principal component of the changes for 1-,3-, 6-month and, 1-, 2-, 5-, 10-year OIS rates in the press release window; (4) changes in the nowcast of real GDP growth ($\Delta GDP\ Nowcast$) and inflation rate ($\Delta HICP\ Nowcast$) contained in ECB staff projections. The change of MRO, forward guidance, and policy surprises are policy controls announced during the press release window. Including these controls rules out the possibility that the influence of uncertain talking is merely a reflection of policy actions or policymakers' economic outlook.

To ensure that the information is disseminated from the ECB's press conference to the mass media but not the opposite direction, we control for the most recent change in the economic uncertainty index (ΔEPU). Initially developed in Baker et al. (2016), this index reflects the policy-related economic uncertainty discussed in ten influential newspapers published in Europe. Since the time when the ECB holds the press conference is usually also the time when the US Department of Labor releases the initial jobless claim, we follow Ehrmann and Fratzscher (2007) to control for jobless claim surprises ($US\ JC\ Surprise$) to exclude potential confounding effect.

We also include controls determined before the press conference day. We control for the stock market momentum (or reversal) by including the daily return and volatility of the previous trading day (*Lagged Ret* and *Lagged Vol*) as control variables. Finally, we introduce two dummy variables to control for the speech styles of the three presidents of the ECB (p=*Draghi* and p=*Duisenberg*).

When reporting the estimation result, the Newey–West covariance estimator is used to produce consistent standard errors in the presence of heteroskedasticity and auto-correlation (Newey and West, 1987).

6 Empirical Results

6.1 Baseline

In the baseline specification, we regress the SX5E return in the press conference window on the first differences of the uncertain talking measures for the Introductory Statement, journalists' questions,

takes place several days after the announcement. However, having anticipated the change, it is reasonable to assume that investors will respond promptly to the anticipated rate change in the financial markets.

and president's answers, without adding control variables.

Table III displays the results. As shown in Columns (1) to (3), when the three uncertain talking measures are included in the regression separately, only the change in uncertain talking measure in the presidents' answers significantly affects the SX5E returns, but not the uncertain talking measure in the Introductory Statements or journalist' questions. This finding remains consistent when we add the three uncertain talking measures in succession displayed in Columns (3) to (5).

[Table III about here]

The estimated marginal effects of $\Delta UncA\%$ are all positive, meaning that the SX5E returns increase when the ECB president uses more uncertain words to answer questions from the journalists. The magnitude is in the range of 0.199 to 0.212, suggesting that a one percent increase in the unexpected uncertain talking measure in answers raises the SX5E returns by approximately 20 basis points on average. Equivalently, when the president says one more uncertain word, the SX5E return will rise by approximately 0.8 basis points on average. This impact of uncertain talking is by no means negligible compared with the effects of more standard monetary policy shocks that are constructed in, for example, Altavilla et al. (2019), on Eurozone stock markets.

The results above show that only the uncertain talking in the ECB president's answers significantly impacts the stock market. It implies that the uncertain talking in the answers brings new information that is not contained in the Introductory Statement or journalists' questions. It further suggests that investors are more attentive to the perceived level of uncertain talking in the answers than in the intended level of uncertain talking in the Introductory Statement. The reason could be that the Introductory Statement is well-prepared by the ECB staff in advance. Therefore the market does not take the uncertain talking in this statement as a signal revealing policymakers' actual considerations on current situations and their policy reactions. In what follows, we concentrate on the impact of uncertain talking in answers and treat uncertain talking measures in other parts as control variables.

The above regressions results have shown a significant positive impact of the uncertain talking in answers ($\Delta UncA\%$) on the SX5E return. To confirm this finding, we include a variety of control variables that are classified into five groups:

- information determined before the press conference day

- textual features of the transcript
- the ECB's policy decisions made shortly before the press conference
- the ECB's staff macroeconomic projections announced at the press conference
- other important news released during the press conference

Table IV reports the estimation results under different specifications in which we sequentially add control variables to the regression equation. The first group of control variables are listed at the bottom of the table and appear in all specifications. The next four groups of control variables are added sequentially. The results in the first column represent the case when we add only the first group of control variables, and the results in the last column correspond to the case when we add all the control variables.

[Table IV about here]

The effect of the change in the uncertain talking measure in answers remains statistically significant at the 1% level. The results from Column (1) suggest that the effect of uncertain talking cannot be explained by the economic policy uncertainty measured from newspapers, the stock market momentum, or presidents' individual talking styles. The results from Columns (2) to (4) imply that the uncertain talking in answers has an independent effect on the stock market that is unrelated to the effects of other textual features. Finally, the results from Columns (5) to (8) shows that the effect of uncertain talking is not affected by other information released before or at the press conference, including monetary policy decisions, macroeconomic projections, and the US initial jobless claim. The signs remain positive in all these specifications, and the sizes stay in the range of 0.200 to 0.237.

Several control variables also have significant effects on the SX5E returns. Consistent with Schmeling and Wagner (2019), the negative relationship between SX5E returns and the negativity measure of the transcript also stands out in our sample. The forward guidance statement significantly increases stock returns, which is reasonable, to the extent that keeping the nominal rate lower for longer is good news to stock markets. On the contrary, neither the change in the MRO rate nor the change in the market-based monetary policy surprises can significantly affect stock returns. This confirms the findings in Altavilla et al. (2019), as there might exist two competing effects of the monetary policy shock on stock returns that cancel each other. The president dummies are significant in some specifications,

suggesting that the average level of returns is different under different presidencies.

To summarize, we discover a significant positive effect of the uncertain talking measure on the SX5E returns within the time window around the ECB's press conference, which cannot be explained by a number of potential confounding factors that are determined before or on the press conference day.

6.2 Daily Returns

We also examine the impact of the uncertain talking measure on a broader range of stock market indices in the daily close-to-close window. All the control variables listed in Table IV are included in the regression.

The results for leading Eurozone stock market indices are displayed in the first panel of Table V. Column (2) shows that the slope coefficient remains significant and positive for the daily SX5E returns, and its size (0.476) is more than twice of that in the press conference window (0.230). This finding suggests that uncertain talking has a sustained effect on the stock market beyond the press conference window, as it might take time for the market to fully digest the message contained in the uncertain talking from the ECB president. Columns (3) to (8) show that the significant positive effect of uncertain talking also exists for other leading stock market indices of the Eurozone. Comparing the magnitude of the effects suggests that uncertain talking has a more substantial influence on indices that comprise large-size firms than indices covering medium- to small-size firms. Besides, uncertain talking has a greater impact on the value index than on the growth index. These results corroborate the findings in Nofsinger (2001) that the trading of large firms is more affected by macroeconomic announcements than the trading of small firms.

[Table V about here]

In the appendix, we show that the effect not only applies to aggregate stock indices, but also a range of country-specific and industrial-specific indices of the Eurozone. Moreover, the effect of the uncertain talking measure on the daily SX5E returns remains significant and positive for more than a week after the press conference. Furthermore, we regress the daily returns of the leading Eurozone stock market indices on all trading days from 2000-2018, to rule out the possibility that the results are driven by a systematic shift in the stock returns on the ECB's press conference day.

6.3 Further Robustness Checks

The above results imply a positive impact of uncertain talking that is significant in both statistical and economic senses. We further substantiate this finding by employing a series of robustness checks in the appendix.

First, we show that the results are robust to the removal of outliers. We show that the effect stays positive at the 1% significance level when observations with the largest influences on the estimate of the slope coefficient are removed from the sample.

We also show that the effect of uncertain talking remains almost unchanged when we use different variables to represent ECB's monetary policies, its macroeconomic projections, or the readability of the transcript, which are all control variables in the regression.

To check whether the effect is sensitive to the specific construction of the uncertain talking measure, we regress the SX5E returns on variants of the original uncertain talking measure. First, we consider two more choices of the scaling parameter in the formula that defines the uncertain talking measure. Second, we consider two alternative assumptions for obtaining the surprised component of the uncertain talking measure. We treat the first difference of the uncertain talking measure as the surprised component before, based on the assumption that investors expect the amount of uncertain talking in the current press conference should be the same as the last one. However, if investors have no expectation of the amount of uncertain talking before the press conference, then the surprised component is simply the level of the measure. At another extreme, if investors use a sophisticated forecasting regression to predict the amount of uncertain talking, then the surprised component becomes the residual of that forecasting regression. We find that the effect of uncertain talking is positive and significant across all different kinds of measures mentioned above.

7 Decomposing the Uncertain Talking

To rationalize the positive and significant response of the stock returns, we need to understand what message is conveyed through uncertain talking. Although all the words from the uncertain word list "*denot[e] uncertainty, with emphasis on the general notion of imprecision*" (Loughran and McDonald, 2011), their meanings and usages are not identical. To solve this issue, we identify different types of

uncertain talking and examines which type significantly drives the stock market.

Instead of manually picking and grouping words, we rely on the latest developments in the fields of natural language processing (NLP) and machine learning (ML) to perform the task. Our method has two advantages. First, manual classification is based on an individual's own knowledge, so different people can give different word groups. Therefore, the identification is not unique, and the meaning of each word group is ambiguous. In comparison, the objectivity and consistency of the NLP and ML algorithms ensure that the resulting word groups are unique with distinguishable meanings. Second, we will show that the NLP and ML algorithms take the contexts of the ECB statements into account when forming word groups. Thus, these word groups precisely reflect the similarity and difference of how these words are used in the ECB statements. On the contrary, a categorization based on subjective judgment might be too general and not appropriate for analyzing the ECB's press conference transcript.

7.1 Quantifying Word Meanings

To begin with, we quantify the meaning of each word in the ECB's statements with a numeric vector (i.e., a *word vector*). Proposed by Mikolov et al. (2013), the idea of embedding word meanings in numeric vectors has revolutionized the NLP research, and it has immensely enhanced the performance of artificial intelligence.

The method that we use to obtain word vectors is based on the distribution hypothesis in linguistic research, which assumes that words that are close to each other tend to have similar meanings (Harris, 1954; Firth, 1957). The generation of word vectors comprises three steps: First, we need to collect a massive amount of textual data. It provides information about which words are close (or distant). Then we use a random numeric vector to represent each word in the data. Finally, we repeatedly update these vectors, so as to increase (or decrease) the inner products of word vectors, if the corresponding words are close (or distant). The iteration stops if these vectors converge under a certain criterion[10].

We use a wide range of documents on the ECB's website as textual data to obtain word vectors. It includes press conference transcripts, policy statements, policy accounts, speeches, interviews, and in-depth reports and analyses. Consequently, we use approximately 440,000 sentences from more

[10] A detailed non-technical introduction to the word vector can be found in the appendix.

than 4,000 documents to obtain the vectors of around 40,000 words. Each 1-by-100 word vector embeds both the semantic and syntactic meanings of the corresponding word.

Since we use only the ECB's documents to generate these vectors, they also contain information about the language style of the ECB's policymakers. This is essential for us to accurately understand the meanings of uncertain words that appear in the ECB's press conference transcripts.

7.2 Identifying Different Types of Uncertain Talking

As the word vector is obtained by maximizing (or minimizing) the inner products of nearby (or distant) words, it is natural to use the cosine distance of two word vectors to quantify the similarity between the two corresponding words. The cosine distance is the normalized inner product of the vectors. The higher the cosine distance is, the closer the meanings of the two words are. Therefore, to classify uncertain words into groups based on their meanings is equivalent to gathering uncertain word vectors with high cosine similarities together.

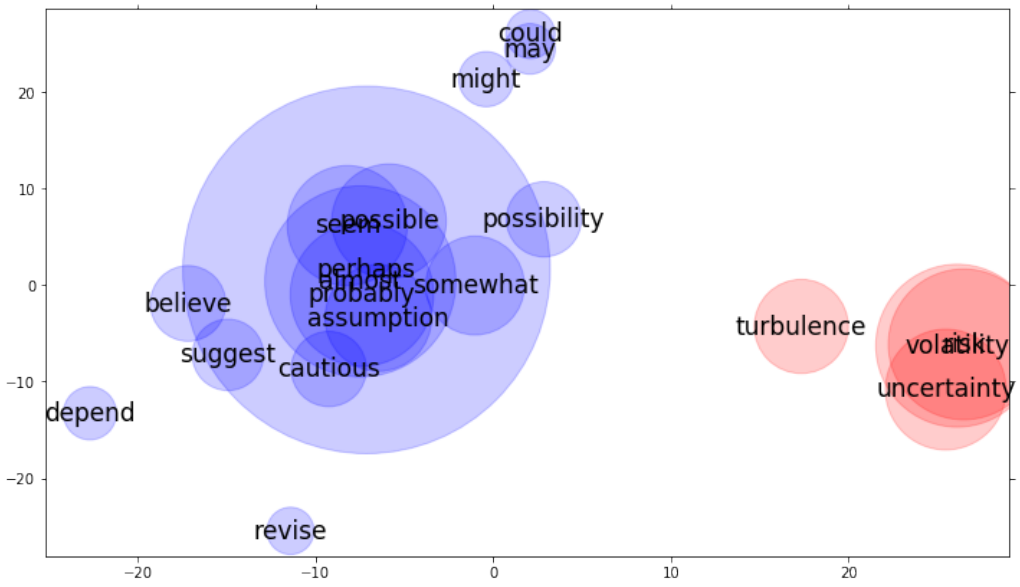
We apply the K-Means clustering algorithm to group word vectors, which partitions N vectors into K clusters by minimizing the error sum of squares within each cluster. As a starting point, we select the first twenty (out of 109) most-mentioned uncertain words and categorize their corresponding word vectors into two clusters for two reasons. First and foremost, assigning the first twenty words into two groups yields the most consistent outcomes, no matter what initial condition we start from in the iterations for generating and clustering word vectors. Besides, these twenty words already take up more than 70% of the total count of uncertain words. As there are various forms of the same words in the dictionary (e.g., risk, risky, riskiness), choosing the first twenty words is sufficient to represent the meaning and usage of all the uncertain words in the press conference transcript.

The two word-clusters derived from the K-Means algorithm are presented in Table VI. As a by-product, we also obtain two new vectors as the two cluster centers, which embed the average meaning of the uncertain words in the two clusters. Accordingly, for a word with the lowest cosine distance with its center vector, its meaning should be most representative for the average meaning of all the words in that cluster. We name clusters after the words with this property, which are *Risk* and *Perhaps* for the two clusters, respectively.

[Table VI about here]

Figure 3
Visualization of Word Vectors and Word Clusters

The figure visualizes the word vectors and word clusters of the top 20 most-mentioned uncertain words. To visualize these vectors, we adopt the Principal Component Analysis (PCA) method to reduce the dimension of word vectors to two. Words from the cluster *Risk* are surrounded by red circles on the right, and words from the cluster *Perhaps* are surrounded by blue circles on the left. The size of the circle signifies the distance from each word to its corresponding cluster center. A larger circle around a word indicates that its word vector is closer to the cluster center, which suggests that the meaning of that word is more similar to the average meaning of its cluster.



Words from the cluster *Risk* account for 24.36% of the total count of uncertain words in the transcripts. From a linguistic point of view, these words denote an undesirable and disorderly situation associated with a potential loss in the future. Sentences with these words are usually used to describe the uncertain and unfavorable economic and financial environments that the ECB faces. Words from the cluster *Perhaps* account for 47.77% of the total count of uncertain words. From a linguistic perspective, these words are related to *hedging language*, which speakers use to limit their commitment to a proposition (Eisenstein, 2019). These words are used by the ECB president to express an uncertain attitude or to deliver a state-dependent statement. Examples of sentences containing words from the two clusters can be found in Table VII.

To gain a direct impression of the classification outcomes, we apply the principal component analysis (PCA) algorithm to reduce the dimensionality of the word vectors and visualize them in the two-dimensional diagram in Figure 3.

[Table VII about here]

Words from clusters *Risk* and *Perhaps* are surrounded by red and blue circles, respectively. The

distance between words from different clusters is larger than the distance between words in the same cluster, indicating that words with the same color indeed have closer meanings. This observation coincides with the classification results derived from NLP and ML algorithms, which demonstrates the strength of such methods in capturing key features of the human language. A larger (smaller) circle suggests that the corresponding word vector is more close (distant) to the cluster center. As discussed above, the largest red and blue circles correspond to the words *Risk* and *Perhaps*, respectively.

We complete the classification by adding other less-mentioned uncertain words to the two clusters. An uncertain word (vector) is included in the cluster *Risk*, if its cosine distance to the center of the cluster *Risk* is shorter than that the center of the cluster *Perhaps*, and vice versa. As a result, words such as "destabilizing," "anomaly," and "fluctuating", which have similar meanings to "risk", "uncertainty", and "turbulence", are categorized in the cluster *Risk*. In contrast, words added to the cluster Perhaps, such as "conditional", "approximately", and "presume", have closer meanings to words "perhaps", "almost", and "probably". The complete word clusters are displayed in Table VIII.

[Table VIII about here]

We then construct two independent uncertain talking measures based on the two word-clusters. The numerator for each measure is now the number of words from the corresponding cluster. The denominator is still the total number of words of the whole transcript. As a result, we differentiate between the environmental uncertain talking and attitudinal uncertain talking. The environmental uncertain talking ($EnvUncK\%_t$) is identified through the cluster *Risk* and is linked to uncertain and unfavorable environments faced by the ECB. The attitudinal uncertain talking ($AttUncK\%_t$) is identified through the cluster *Perhaps* and mostly reflect the uncertain attitude of the ECB policymakers. In the appendix, we further show that only the environmental uncertain talking measure increases with variables that indicate uncertain and unfavorable economic and financial environments of the Euro area.

8 Explaining the Positive Reaction of Stock Returns

To inspect the positive reaction of the SX5E returns, we study the impacts of the environmental and attitudinal uncertain talking measures on the SX5E returns in the press conference window.

As shown in Table IX, only the change of the environmental uncertain talking significantly increases the SX5E returns, regardless of whether or not control variables are added to the regression equation. It means that stock price increases when the ECB president uses more words such as "risk", "uncertainty", and "turbulence" to answer the journalists" questions. Since the impact of the attitudinal uncertain talking is not significant and the impact of the environmental uncertain talking is even stronger than the impact of the overall measure, it is reasonable to conclude that the environmental uncertain talking is the major cause behind the positive SX5E returns. In what follows, we aim to discover why the environmental uncertain talking raises the stock returns.

[Table IX about here]

8.1 An Illustrative Example

To comprehend the finding above, consider the ECB's press conference held on October 22nd, 2015. The monetary policy decision released before the press conference remains the same as the previous one. However, during the press conference, President Draghi uses more words than the last time from the cluster *Risk* when answering questions. In response to a question on further rate cuts, he thoroughly discusses the *risks* from weakened demand and subdued inflation expectations. In another answer, he acknowledges the political *uncertainty* that prevailed in some European countries during that period.

The SX5E returns move up 2.04% during this press conference window. The next day, a news report entitled "ECB Signals Stimulus Boost" appears in the Wall Street Journal. This report also explicitly cites President Draghi's answers that contain words from the cluster *Risk* and links them to the ECB's dovish inclination:

ECB President Mario Draghi signaled Thursday that the bank is prepared to undertake another large stimulus package . . . , as officials struggle with ultra-low inflation and a tepid recovery. . . . "These risks have gone up and we want to be vigilant," he said.

As is implied in this example, it does seem that the ECB president shows a more uncertain and unfavorable condition by using more words from the cluster *Risk*. At first glance, this might bring pessimistic sentiment to the markets. However, the above report suggests a different story. Investors interpret the president's judgment of a worsened environment as a signal of more expansionary mon-

etary policy moves in the future, which itself is good news to the markets.

This example suggests that the environmental uncertain talking affects the stock market through two channels that work against each other. The uncertain and unfavorable environment revealed by the environmental uncertain talking is negative news that induces lower stock returns. Nevertheless, if investors believe that the monetary policy is state-dependent, then the adverse environment also implies more policy accommodations in the future, which is positive news that raises stock returns. Equity prices will likely increase when the latter effect outweighs the former one.

8.2 The Two-Channel Interpretation

.The example discussed above indicates that the environmental uncertain talking might affect stock returns through two mechanisms. To follow the current literature, we refer to them as the environmental channel and the policy channel.

To explain how these two channels work, we first introduce a conceptual framework to analyze the effect of the central bank's policies on stock returns, which is also employed by Bernanke and Kuttner (2005) and Cieslak and Schrimpf (2019) for a similar purpose. In general, three forces can influence stock returns: changes in the expected dividend growth, changes in the expected real interest rate, and changes in the expected excess returns (that is, the risk premia). Based on this idea, Campbell (1991) and Campbell and Ammer (1993) apply a log-linear approximation method to decompose the stock returns into three components in the following dynamic accounting identity:

$$e_t^y = \tilde{e}_t^d - \tilde{e}_t^r - \tilde{e}_t^{ex} \tag{5}$$

e_t^y is the unexpected stock return at time t. $\tilde{e}_t^d$ is the revision of expectations of discounted future dividend growth. The positive sign suggests that, holding the other two factors constant, if the expected dividend growth increases, the unexpected return should be positive. $\tilde{e}_t^r$ is the revision of expectations on discounted future real interest rates. The negative sign implies that the stock price decreases with higher expectations of real interest rates. $\tilde{e}_t^{ex}$ is the expectations revision on discounted future risk premia. If equity investors expect higher risk premia in the future, then the unexpected stock return should be negative to compensate for the risk. This identity shows that stock market returns change with, and only with, new information influencing factors on the right side of the equation.

The environmental channel works when investors update their expectations of economic and financial conditions based on the new information from the central bank. This channel captures the direct effect of the environmental uncertain talking, as the ECB president uses words from the cluster *Risk* to describe uncertain and unfavorable situations. If an increased environmental uncertain talking makes investors believe that there are more risks and uncertainties than expected, then an upward shift in the expected risk premia ($\tilde{e}_t^{ex} > 0$) will lower the stock returns. Besides, if investors feel more pessimistic about future states of the economy, there will be a downward adjustment in their expectations of future dividend growth ($\tilde{e}_t^{d} < 0$), which also reduces stock returns.

The policy channel works when investors update their expectations of future monetary policy decisions based on the new information from the central bank. This channel captures the indirect effect of uncertain talking. If investors believe that the monetary policy made by the policymakers of the ECB is contingent on their views on economic and financial environments, an increase in the environmental uncertain talking might then be understood as a signal of more expansionary monetary policies in the future. On the one hand, a more expansionary policy implies lower future discount rates ($\tilde{e}_t^{r} < 0$), which results in higher stock returns. On the other hand, more expansionary policy is connected with lower expected risk premia ($\tilde{e}_t^{ex} < 0$) because it can reduce the riskiness of stocks by lowering firms' financing costs (Bernanke and Kuttner, 2005), which also raises stock returns. The stock price should increase because of these two forces.

In summary, the environmental uncertain talking affects stock returns through two channels in the opposite direction. It can lead to lower stock returns through the environmental channel but also result in higher stock returns through the policy channel. The overall positive sign in the regression results implies that the policy channel plays a dominant role.

One possible explanation is that, while the central bank announcement is merely one of many sources from which market participants can learn about economic and financial environments, it is the only source from which they learn about *policymakers' views* on these conditions. These views might not be extremely helpful in discovering the true state of the economy (Bauer and Swanson, 2020), but they are essential to uncover the future monetary policy stance. Thus, when the ECB president reveals his or her actual thinking about the environments by using uncertain words, investors are more inclined to revise their expectations of monetary policy through the policy channel, rather than their expectations of those environments through the environmental channel. This explains why the policy

channel prevails.

9 Evidence for the Two-Channel Interpretation

In this section, we present evidence to support the argument that the positive reaction of stock returns results from the dominance of the policy channel. We first confirm the existence of the policy channel by showing that the environmental uncertain talking influences the expectations of ECB's future monetary policy stance. We then compare the numerical and textual signals that the ECB releases (i.e., "hard information" and "soft information"). We show that the environmental channel is relatively weakened when the ECB releases numeric signals about its economic projections on press conference day. Similarly, the policy channel is relatively weakened when the ECB releases numeric signals about its monetary policy.

9.1 The Existence of the Policy Channel

To begin with, we explore whether the environmental uncertain talking affects the Euro Overnight Index Average (EONIA)-based overnight indexed swap (OIS) rates[11]. As the current overnight benchmark interest rate for the Euro, the EONIA rates closely follow the official monetary policy rate set by the ECB. The fixed-rate in this swap contract is referred to as the EONIA-based OIS rate, which reflects market-based expectations of the ECB's policy rate decided during the period in which the contract is being executed. A growing number of studies (e.g., Altavilla et al., 2019) use changes in the OIS rates in the narrow time window around the ECB's policy events to gauge the impact of its monetary and communication policies.

We consider the OIS rates of nine maturities from six months up to twenty years. For each OIS rate, we regress its percentage change in the press conference window on the changes in the uncertain talking measures. If the environmental uncertain talking can significantly revise market participants' beliefs about the future policy stance through the policy channel, then an increase in the environmental uncertain talking measure should significantly lower the OIS rates in the press conference window.

The first panel of Table X displays the results. The environmental uncertain talking has a significant negative impact on the OIS rates indeed. The effects are negative, meaning that when the ECB

[11] In the EONIA-based OIS contract, one party swaps a floating interest rate cash flow in exchange for a fixed interest rate cash flow from another party, in which the EONIA rate serves as the benchmark floating rate.

president uses more words from the cluster *Risk* to answer questions, it will cause downward revisions in market expectations of future monetary policy rates. This result confirms the existence of the policy channel.

[Table X about here]

Comparing the estimated coefficients for a spectrum of the OIS rates at various maturities shows that the magnitude of the effect first becomes larger as maturity increases. It peaks at the three-year maturity and then gradually decreases. Thus, the environmental uncertain talking mainly affects the medium- to long-term expectations of the ECB's policy stance.

We further check the validity of the two-channel interpretation by inspecting the reaction of the German government bond yields in the press conference window. Similar to the OIS rates, the nominal German government bond yields have also been used to reflect market-based expectations of monetary policy rates in several articles, including Altavilla et al. (2019) and Kerssenfischer (2019).

Though the changes in OIS rates are highly correlated with the changes in German government bond yields in the press conference window, these two assets are different in many aspects. Regarding our analysis, while uncertain talking affects the OIS rates primarily through the policy channel, the bond yields might be influenced through both the environmental and policy channels. As a risk-free asset in the Eurozone, when market participants feel the economic and financial conditions are more uncertain and turbulent, their demand for the German government bond increase and therefore bond yields decrease. When market participants expect more expansionary monetary policy moves, nominal bond yields also decrease. As a result, an increase in the environmental uncertain talking is expected to reduce bond yields, no matter whether through the environmental or policy channel.

The second panel of Table X confirms our two-channel interpretation. The negative influences of the environmental uncertain talking on nominal bond yields are significant at 10% level for maturities spanning from one to ten years and at 5% level for maturities spanning from one to four years.

It should be noted that neither the overall uncertain talking measure nor the attitudinal uncertain talking measure has a significant effect on the OIS rates or nominal government bond yields. This matches our findings on the stock market. The discussion of these results can be found in the appendix.

9.2 The Effects of Hard and Soft Information

In this section, we discuss the relative importance of the policy and environmental channels in subsamples, which are formed based on whether or not the ECB releases hard information about its monetary policies and macroeconomic projections on press conference day.

The ECB discloses two types of information on press conference day: Hard information represented as numbers and soft information represented as texts. It always releases soft information through presidents' talks on press conference day. However, it only releases hard information about its policy when there is a change in the conventional or unconventional measure of monetary policy. It only releases hard information about its view on the economic conditions when the staff macroeconomic projection is announced once in a quarter.

Although these two types of information can contain signals conveying similar meanings, the signal communicated through hard information as numbers is more direct, which is easier for people to focus on and form consistent expectations from. According to the Bayes' rule, if both signals are available, agents should place more weight on the one with higher precision, which is the hard information, when updating beliefs. As a result, if the ECB reveals its view on the current environment through numeric projections, then the environmental channel of the environmental uncertain talking is expected to be weakened. Similarly, if the ECB announces a quantitative change of its policy rates, then the policy channel of the environmental uncertain talking is expected to be weakened.

We first split the samples into days with and without a publication of the ECB staff macroeconomic projection. The projection is updated in the last month of each quarter, and it contains the forecasts of main macroeconomic variables provided by the staff of the ECB or Eurosystem. It was published a few days after the press conference before the second quarter of 2004. After that, it is released on press conference day: An advanced snapshot is announced in the Introductory Statement at the press conference, and then the full projection is published after the press conference.

When the projection is announced on press conference day, investors can learn about ECB's view on economic and financial conditions through both numeric projections and uncertain talking. Thus, the uncertain talking becomes less valuable for inferring these conditions. If so, the environmental channel will become weakened relative to the policy channel. The two-channel interpretation predicts that the positive effect of uncertain talking will be stronger and more significant on press conference

days with a projection than on days without one.

The regression results displayed in Panel (a) of Table XI confirms the prediction. While the estimated slope coefficient for the environmental uncertain talking in the first sub-sample (conference days without a projection) is not significant even on 10% level, it remains significant at 5% or lower level in the second sub-sample (conference days with a projection). The size of the coefficient (0.202) in the first sub-sample is less than half of the size of the coefficient in the second one (0.436). There is also a clear difference in the R^2 of the two regressions.

[Table XI about here]

We then split samples into days with quantitative changes in monetary policies and days without a policy change. A change in the monetary policy means a change in one of the ECB's three monetary policy rates, a change in the conduct of the forward guidance policy, or a change in the ECB's asset purchasing program[12].

Similar to the discussion above, when a policy change occurs, investors have access to two signals to learn about the ECB's monetary policy. One is the hard information as the quantitative change, and the other is the soft information as the uncertain talking[13]. Again, the Bayes' rule depicts that market participants place more weight on hard actions than soft words. As a consequence, the policy channel of the environmental uncertain talking will become weakened relative to the environmental channel. The two-channel interpretation predicts that the positive effect of uncertain talking will be weaker and less significant on press conference days with a policy changes than on days without one.

The results displayed in Panel (b) of Table XI confirms this prediction. There are even more substantial differences in the estimated slope coefficient for the environmental uncertain talking between the two sub-samples in Panel (b) than those in Panel (a). The coefficient is significant at 1% level for the

[12]In our analysis, we define a policy change from the perspective of the central bank, that is whether a change exists in ECB's monetary policy. Alternatively, one might feel more appropriate to measure the change from the perspective of the financial market and use the change in the OIS rates to gauge the monetary policy surprise. To follow this idea, we ran regressions on sub-samples with small and large (absolute) changes in the OIS rates but find no noticeable difference. It implies that the market might pay different attention to the president's talk between cases with or without a change in the policy, but not between cases with or without a sizable monetary policy surprise. Besides, there are other disputable issues associated with the market-based monetary policy surprise. For example, there is no clear evidence on whether one should focus on changes in the short-end of the OIS rates, the long-end of the OIS rates, or whether the monetary policy shock should be extracted from the OIS rates or bond yields.

[13]The forward guidance policy can also be regarded as a verbal signal. However, in our specification, we do not consider the details of the forward guidance statement but only consider a change when the ECB starts to include or remove such a statement in its decision.

first sub-sample (conference days without policy action), but the same coefficient is not significant at any standard significance levels for the second sub-sample (conference days with policy action). Furthermore, while the effect of uncertain talking remains positive in the first sub-sample, it turns into negative in the second sub-sample. This is a convincing piece of evidence that corroborates the two-channel interpretation because it suggests that, when a policy change exists, the policy channel becomes so weak that it is even dominated by the environmental channel, which causes a reversal in the direction of the overall effect. The R^2 in the first sub-sample is much lower than that in the second sub-sample, which might be explained by the sizeable difference in the number of observations in two sub-samples.

10 Conclusion

Is uncertain talking in the ECB's press conferences news or noise to the stock markets? We answer this question by quantifying the amount of uncertain talking at the ECB press conferences and showing that it significantly raises the stock returns of the Eurozone.

This finding is extremely robust to various alternative specifications, but it is also at odds with the predictions from the existing literature. To explain this result, we decompose the overall uncertain talking into two dimensions and show that only the environmental uncertain talking, which is empirically related to uncertain and unfavorable economic and financial environments that the ECB faces, can significantly drive the markets.

The effect is interpreted through the environmental and policy channels. The environmental channel works when investors consider the environmental uncertain talking as a negative signal of economic and financial conditions, which lowers stock returns. In contrast, in the policy channel, increased environmental uncertain talking is a positive signal of future expansionary monetary policy, which increases stock returns. The overall positive effect implies that the policy channel dominates. We find further empirical evidence to substantiate this interpretation on the bond market and interest swap market, and by comparing the relative importance of hard and soft information in different sub-samples.

Our study lays foundations for future research and policy-making. First, we propose a new approach to objectively quantify word meanings based on the context. Just as Loughran and McDonald (2011)

pointed out, words in a different context can be used differently. Our approach can be applied in other research to automatically derive accurate measures of certain concept by grouping words with similar meanings together in textual data. Second, we provide an empirical foundation for modeling the relationship between the provision of public information from policymakers and the reaction of asset prices. Since none of the existing literature matches our empirical evidence, modeling how policymakers' assessment of economic conditions affects market participants' expectations of future policy paths will be a fruitful direction for future research.

References

Ai, H. (2010). Information quality and long-run risk: asset pricing implications. *The Journal of Finance*, 65(4):1333–1367.

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

Andrade, P., Gaballo, G., Mengus, E., and Mojon, B. (2019). Forward guidance and heterogeneous beliefs. *American Economic Journal: Macroeconomics*, 11(3):1–29.

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

Bauer, M. and Swanson, E. (2020). The fed's response to economic news explains the "fed information effect". *Manuscript, University of California Irvine*.

Bernanke, B. (2007). Monetary policy under uncertainty. *The 32nd Annual Economic Policy Conference, Federal Reserve Bank of St. Louis*.

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

Born, B., Ehrmann, M., and Fratzscher, M. (2014). Central bank communication on financial stability. *The Economic Journal*, 124(577):701–734.

Brevik, F. and d'Addona, S. (2010). Information quality and stock returns revisited. *Journal of Financial and Quantitative Analysis*, 45(6):1419–1446.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The economic journal*, 101(405):157–179.

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

Cieslak, A., Morse, A., and Vissing-Jorgensen, A. (2019). Stock returns over the fomc cycle. *Journal of Finance, forthcoming*.

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

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

Cohen, L., Malloy, C., and Nguyen, Q. (2020). Lazy prices. *The Journal of Finance*, 75(3):1371–1415.

Dzieliński, M., Wagner, A., and Zeckhauser, R. (2017). Straight talkers and vague talkers: The effects of managerial style in earnings conference calls. Technical report, National Bureau of Economic Research.

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

Eisenstein, J. (2019). Natural language processing. Manuscript.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., and Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, 33(3):1184–1216.

Epstein, L. and Schneider, M. (2008). Ambiguity, information quality, and asset pricing. *The Journal of Finance*, 63(1):197–228.

Firth, J. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in Linguistic Analysis*.

Goldberg, Y. (2017). Neural network methods for natural language processing. *Synthesis Lectures on Human Language Technologies*, 10(1):1–309.

Gorodnichenko, Y. and Weber, M. (2016). Are sticky prices costly? evidence from the stock market. *American Economic Review*, 106(1):165–199.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99(S1):114–133.

Harris, Z. (1954). Distributional structure. *Word*, 10(2-3):146–162.

Hautsch, N. and Hess, D. (2007). Bayesian learning in financial markets: Testing for the relevance of information precision in price discovery. *Journal of Financial and Quantitative Analysis*, 42(1):189–208.

Jegadeesh, N. and Wu, D. (2017). Deciphering fedspeak: The information content of fomc meetings. Manuscript.

Kerssenfischer, M. (2019). Information effects of euro area monetary policy: New evidence from high-frequency futures data. *Deutsche Bundesbank Discussion Paper Series*, (7).

Kostovetsky, L. and Warner, J. (2020). Measuring innovation and product differentiation: Evidence from mutual funds. *The Journal of Finance*, 75(2):779–823.

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

Li, G. (2005). Information quality, learning, and stock market returns. *Journal of Financial and Quantitative Analysis*, 40(3):595–620.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2013). Ipo first-day returns, offer price revisions, volatility, and form s-1 language. *Journal of Financial Economics*, 109(2):307–326.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

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

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., and McClosky, D. (2014). The stanford corenlp natural language processing toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

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

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Nofsinger, J. (2001). The impact of public information on investors. *Journal of Banking & Finance*, 25(7):1339–1366.

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

Schmeling, M. and Wagner, C. (2019). Does central bank tone move asset prices? Manuscript.

Tetlock, P. (2014). Information transmission in finance. *Annual Review of Financial Economics*, 6(1):365–384.

Tobback, E., Nardelli, S., and Mertens, D. (2017). Between hawks and doves: Measuring central bank communication. Technical report, European Central Bank.

Ulrich, M., Jakobs, E., and Tran, R. (2019). The real-time impact of ecb press conferences on financial markets. *Available at SSRN 3436852*.

Table II
Summary Statistics

This table reports the summary statistics for the textual features of the ECB's press conference transcripts in Panel (a) and the stock returns at press conference day in Panel (b). The sample period is from Dec 2000 to Sep 2018. In Panel (a), *UncA%*, *UncN%* and *UncQ%* are the uncertain talking measure in the presidents' answer, the Introductory Statement and the questions from journalists, respectively. *NegA%*, *NegI%* and *NegQ%* are the respective negativity measures. *TWords* is the total number of words in the transcript. The five columns on the left side contain the summary statistics for their levels, and the five columns on the right side contain the summary statistics for their first differences. Panel (b) shows stock returns in the press conference window and in the daily close-to-close window in percentage terms. The index in the press conference window is SX5E, namely the Euro STOXX 50 index. The indices in the daily window are the SX5E, Europe STOXX Large 200, Europe STOXX Median 200, Europe STOXX Small 200, MSCI Europe, MSCI Europe Growth, and MSCI Europe Value indices.

(a) Textual Features

	Level					Difference				
	Mean	Std.Dv	Min	Max	No.Obs	Mean	Std.Dv	Min	Max	No.Obs
UncA%	1.504	0.443	0.580	3.319	194	0.004	0.507	-1.777	1.312	194
UncN%	0.785	0.285	0.264	1.682	194	-0.000	0.251	-0.689	0.764	194
UncQ%	0.930	0.298	0.326	2.041	194	0.001	0.368	-1.132	1.427	194
NegA%	3.591	0.769	1.875	5.334	194	0.003	0.810	-2.313	2.730	194
NegN%	1.496	0.626	0.444	3.850	194	-0.006	0.523	-2.104	2.460	194
NegQ%	1.947	0.516	1.015	4.296	194	-0.001	0.580	-1.697	2.242	194
TWords	2554.247	596.550	890	3692	194	5.768	515.668	-2463	2218	194

(b) Stock Returns

	Mean	Std.Dv	Min	Max	No.Obs
Press conference window return:					
SX5E	-0.072	0.593	-3.004	2.040	193
Daily returns:					
SX5E	-0.032	1.610	-6.220	5.669	194
STOXX Large 200	-0.031	1.329	-5.617	4.941	194
STOXX Mid 200	0.017	1.272	-5.352	4.594	194
STOXX Small 200	0.058	1.263	-4.964	5.704	194
MSCI	-0.033	1.481	-6.130	5.302	194
MSCI Value	-0.030	1.591	-6.515	6.365	194
MSCI Growth	-0.032	1.390	-5.693	3.862	194

Table III
Reaction of SX5E Returns to Changes in Uncertain Talking Measures (without Controls)

This table reports the results for regressions of SX5E returns on the changes in the uncertain talking measures. The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window, which is defined in Altavilla et al. (2019). $\Delta UncN\%$, $\Delta UncQ\%$ and $\Delta UncA\%$ are changes in the uncertain talking measures in the Introductory Statement, journalists' questions and presidents' answers respectively, calculated based on Equation (1). Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	(1)	(2)	(3)	(4)	(5)
$\Delta UncN\%$	-0.174				-0.086
	(0.159)				(0.164)
$\Delta UncQ\%$		0.114		0.096	0.094
		(0.113)		(0.115)	(0.113)
$\Delta UncA\%$			0.212***	0.208***	0.199***
			(0.066)	(0.068)	(0.072)
Constant	-0.072*	-0.071*	-0.072*	-0.072*	-0.072*
	(0.040)	(0.039)	(0.040)	(0.040)	(0.040)
Observations	193	193	193	193	193
$R^2(\%)$	0.5	0.5	3.3	3.6	3.7

Table IV
Reaction of SX5E Returns to Changes in Uncertain Talking Measures (with Controls)

This table reports the results of regressing SX5E (Euro Stoxx 50) returns on the change of uncertain talking measure in presidents' answers. The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window. $\Delta UncA\%$ is the change in the uncertain talking measure in presidents' answers on the ECB's press conference. We use five sets of control variables: (i) Variables determined prior to press conference day: ΔEPU is the monthly change in the policy uncertainty index; *Lagged ret* and *Lagged vol* are the daily return and realized volatility of the SX5E index on the day before the press conference day; $p = Duisenberg$ and $p = Trichet$ are two dummy variables indicating the presidency. Those controls are included in all columns. (ii) Textual features of the transcript: $\Delta UncN\%$ and $\Delta UncQ\%$ are changes in the uncertain talking measures from the Introductory Statement and journalists' questions; $\Delta NegA\%$, $\Delta NegN\%$ and $\Delta NegQ\%$ are changes in the negativity measures of the answers, statements and questions; $\Delta TWords$ is the change in the total number of words in the transcript. (iii) The ECB's policy decisions: ΔMRO is the change in the marginal refinancing operations rate; FG is a dummy indicating the existence of the forward guidance statement; *PRShock* is the first principal component extract from the changes in the EONIA-based OIS rates in the press-release window, which is regarded as the market-based monetary policy shock. (iv) The ECB's staff macroeconomic projections: $\Delta GDPNowcast$ and $\Delta HICPNowcast$ are the changes in the ECB's staff projections of Euro area's real GDP growth and HICP inflation rate of the current year. (v) Other macroeconomic announcements: *US JC Surprise* is the surprise component of the US initial jobless claims released at the same time when the ECB's press conference takes place. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta UncA\%$	0.213***	0.200***	0.223***	0.237***	0.222***	0.222***	0.230***	0.230***
	(0.067)	(0.072)	(0.081)	(0.083)	(0.076)	(0.075)	(0.078)	(0.078)
$\Delta UncN\%$		-0.095	0.044	0.008	-0.014	-0.010	-0.029	-0.029
		(0.176)	(0.156)	(0.176)	(0.172)	(0.174)	(0.177)	(0.177)
$\Delta UncQ\%$		0.096	-0.025	-0.034	-0.029	-0.025	-0.039	-0.039
		(0.117)	(0.116)	(0.116)	(0.111)	(0.110)	(0.111)	(0.112)
$\Delta NegA\%$			-0.104**	-0.104**	-0.094*	-0.098*	-0.095*	-0.095*
			(0.051)	(0.052)	(0.051)	(0.053)	(0.055)	(0.055)
$\Delta NegN\%$			-0.239***	-0.257***	-0.248***	-0.258***	-0.267***	-0.267***
			(0.071)	(0.079)	(0.083)	(0.083)	(0.083)	(0.083)
$\Delta NegQ\%$			0.129*	0.124*	0.115	0.111	0.117*	0.117
			(0.071)	(0.071)	(0.071)	(0.070)	(0.071)	(0.071)
$\Delta TWords$				-0.000	-0.000	-0.000	-0.000	-0.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ΔMRO					0.016	0.014	-0.158	-0.160
					(0.481)	(0.508)	(0.511)	(0.505)
FG					0.551**	0.550**	0.493**	0.493**
					(0.214)	(0.216)	(0.213)	(0.214)
PRShock						-0.037	-0.029	-0.029
						(0.061)	(0.062)	(0.062)
$\Delta GDPNowcast$							0.071	0.071
							(0.049)	(0.048)
$\Delta HICPNowcast$							0.036	0.036
							(0.060)	(0.061)
US JC Surprise								-0.000
								(0.002)
ΔEPU	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged ret	0.006	0.005	0.016	0.015	0.002	0.003	0.011	0.011
	(0.032)	(0.033)	(0.031)	(0.031)	(0.033)	(0.033)	(0.034)	(0.035)
Lagged vol	-0.002	-0.002	-0.002	-0.002	0.001	0.001	0.004	0.004
	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.011)	(0.011)	(0.011)
p=Duisenberg	0.006	0.009	0.010	0.013	0.391*	0.388*	0.327*	0.327*
	(0.121)	(0.123)	(0.124)	(0.124)	(0.200)	(0.200)	(0.196)	(0.196)
p=Trichet	0.089	0.092	0.086	0.091	0.472**	0.477**	0.423**	0.423**
	(0.099)	(0.099)	(0.099)	(0.100)	(0.193)	(0.195)	(0.192)	(0.193)
Constant	-0.114	-0.115	-0.115	-0.117	-0.499***	-0.501***	-0.459**	-0.459**
	(0.094)	(0.095)	(0.095)	(0.096)	(0.187)	(0.188)	(0.183)	(0.183)
Observations	193	193	193	193	193	193	193	193
$R^2(\%)$	4.0	4.5	10.3	10.7	17.1	17.4	19.1	19.1

Table V
Reaction of Stock Returns to the Change in Uncertain Talking Measures (Daily Returns)

This table reports the results for regressions of stock returns in the close-to-close daily window on the change in the uncertain talking measure in answers. The sample period is from Dec 2000 to Sep 2018. $\Delta UncA\%$ is the change in the uncertain talking measure in presidents' answers on the ECB's press conference. All the control variables listed on Table IV are included in the regressions. The dependent variables are the daily returns of leading stock market indices of the Eurozone, which comprises SX5E (Euro Stoxx 50), Euro STOXX Large 200, Euro STOXX Median 200, Euro STOXX Small 200, MSCI Europe, MSCI Euro Growth and MSCI Euro Value. We also present the reaction of the SX5E return in the intra-day press conference window in Column (1) as benchmarks. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dept. variable	SX5E	SX5E	STOXX Large	STOXX Median	STOXX Small	MSCI	MSCI Growth	MSCI Value
Time window	PC	daily	daily	daily	daily	daily	daily	daily
$\Delta UncA\%$	0.230***	0.476**	0.369**	0.303	0.289*	0.425**	0.351*	0.485**
	(0.078)	(0.201)	(0.179)	(0.184)	(0.171)	(0.187)	(0.186)	(0.197)
Observations	193	194	194	194	194	194	194	194
$R^2(\%)$	19.1	15.5	14.9	14.5	15.1	15.8	16.8	15.2

Table VI
Word Clusters for the 20 Most-Mentioned Uncertain Words

This table displays the two word clusters for the top 20 most-mentioned uncertain words in the ECB's press conference transcript. We first employ an unsupervised neural network to obtain numeric representations of words in ECB's announcements. These word vectors embed the meanings of the corresponding words. We then apply the K-Means algorithm to classify the vectors of the top 20 most mentioned uncertain words into two clusters. Each cluster is named after the word whose vector is closest to the cluster center, which represents average meaning of all the words in that cluster.

Cluster Name	Percentage	Word
Risk	24.36%	risk, uncertainty, volatility, turbulence
Perhaps	47.77%	could, may, believe, might, possible, perhaps, seem, depend, probably, assumption, possibility, suggest, somewhat, almost, cautious, revise

Table VII
Examples of Sentences with Words from the Clusters *Risk* and *Perhaps*

This table provides three examples of sentences that contain words from the cluster *Risk* and three examples of sentences that contain words from the cluster *Perhaps*. These sentences were used by the ECB presidents at press conferences to answer questions in the Q&A session. Please refer to the appendix for an exhaustive list of all the sentences containing these words. Source: ECB.

Cluster	Example
Risk	*The other one is the uncertainty about the economic prospects and the low level of confidence ...* - President Duisenberg, 8 November 2011
	This is part of the overall financial turbulence that we have to cope with. - President Trichet, 15 January 2009
	This is one potential risk of enhanced financial volatility and we 're also monitoring ... - President Draghi, 13 September 2018
Perhaps	*No, it could be a signal that markets have to listen more to me than to others.* - President Duisenberg, 8 November 2011
	It seems to me that, ..., we are in a somewhat different universe to that of ten or even five years ago. - President Trichet, 07 August 2008
	And perhaps also one needs to look at the space, the fiscal space the government has ... - President Draghi, 22 October 2015

Table VIII
Word Clusters for All the Uncertain Words

This table displays the two complete word clusters, which are generated by extending the two clusters described in Table VI to all uncertain words. An uncertain word (vector) is included in the cluster *Risk*, if its cosine distance to the center of the cluster *Risk* is shorter than that the center of the cluster *Perhaps*, and vice versa. We count the words in the clusters *Risk* and *Perhaps* independently, and use them to build up the *environmental* uncertain talking measure and the *attitudinal* uncertain talking measure respectively.

Cluster Name	Percentage	Word
Risk	28.15%	risk, uncertainty, volatility, turbulence, unexpectedly, fluctuate, unhedged, risky, deviation, speculate, unexpected, intangible, riskiest, vagary, speculative, riskier, unidentified, variant, arbitrariness, fluctuation, deviate, incompleteness, variance, risk, uncertain, exposure, unknown, variable, arbitrary, unsettled, volatile, destabilizing, speculation, vary, unusually, instability, precaution, pend, anomaly, suddenly, anomalous, uncertainty, alteration, precautionary, riskiness, improbable, variability, variation, unusual, volatility, sudden
Perhaps	71.85%	could, may, believe, might, possible, perhaps, seem, depend, probably, assumption, possibility, suggest, somewhat, almost, cautious, revise, appear, unobservable, cautiously, ambiguous, apparent, contingency, unpredictable, reconsider, probable, unpredictability, differ, doubtful, approximately, predictive, predictability, dependent, unclear, uncertain, confusion, presume, dependency, confuse, contingent, crossroad, recalculate, likelihood, conceivable, reassess, hinge, conceivably, indefinitely, occasionally, tentatively, imprecise, ambiguity, predictor, possibly, approximate, anticipate, maybe, hide, doubt, nearly, sometime, conditional, presumably, untested, reassessment, dependence, unfamiliar, apparently, probability, reinterpret, sometimes, assume, somewhere, seldom, predict, roughly, clarification, random, tentative, probabilistic, approximation, arbitrarily, presumption, anticipation, preliminary, prediction, crossroads, indefinite, tend, sporadically

Table IX
Reaction of SX5E Returns to Changes in Two Types of Uncertain Talking Measure

This table reports the results for regressions of SX5E returns on the changes in the two types of the uncertain talking measures in answers. The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window. $\Delta EnvUncA\%$ and $\Delta AttUncA\%$ are the first differences of the environmental and attitudinal uncertain talking measures in answers. They are calculated based on Equation (1). Words that we use to build up these measures are listed in Table VIII. Columns (1), (3) and (5) present the results when no control variables is included. Columns (2), (4) and (6) present the results when all the control variables listed in Table IV are included. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta EnvUncA\%$	0.295***	0.291***			0.280***	0.285***
	(0.090)	(0.089)			(0.093)	(0.094)
$\Delta AttUncA\%$			0.149	0.158	0.123	0.149
			(0.102)	(0.116)	(0.104)	(0.115)
Controls	No	All	No	All	No	All
Observations	193	193	193	193	193	193
R^2 (%)	2.8	18.3	0.9	16.9	3.4	19.0

Table X
Reactions of OIS Rates and German Government Bond Yields to Changes in Environmental Uncertain Talking Measure

This table reports the results for regressions of changes in the EONIA-based Overnight Indexed Swap (OIS) rates and nominal German government bond yields on the change in the environmental uncertain talking measure in answers. The sample period is from Dec 2000 to Sep 2018. The dependent variables in the upper panel are the changes of the OIS rates, in the bottom panel are the changes in the German government bond yields, both within the press conference time window. The maturities of the OIS rates and German government bond yields span from six months to 20 years. $\Delta EnvUncA\%$ is the first difference of the environmental uncertain talking measure in answers. It is calculated based on Equation (1). Words that we use to build up this measure are listed in the cluster *Risk* of Table VIII. All the control variables listed on Table IV are included in the regressions. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) Overnight Indexed Swap Rates

	OIS6M	OIS1Y	OIS2Y	OIS3Y	OIS5Y	OIS8Y	OIS10Y	OIS20Y
$\Delta EnvUncA\%$	-0.996	-1.708**	-1.766**	-1.947**	-1.490**	-0.986*	-1.037**	-1.024**
	(0.606)	(0.827)	(0.843)	(0.771)	(0.749)	(0.538)	(0.480)	(0.450)
Observations	193	193	193	193	193	193	193	167
R^2 (%)	10.3	11.6	13.4	14.3	12.9	14.8	15.3	22.4

(b) German Government Bond Yields

	DE6M	DE1Y	DE2Y	DE3Y	DE5Y	DE8Y	DE10Y	DE20Y
$\Delta EnvUncA\%$	-1.111	-2.016**	-2.048**	-1.867**	-1.523*	-1.053*	-1.071*	-0.713
	(0.767)	(0.938)	(0.859)	(0.803)	(0.786)	(0.606)	(0.555)	(0.540)
Observations	134	163	193	193	193	193	193	193
R^2 (%)	9.7	11.4	13.0	12.2	12.8	13.5	13.2	13.9

Table XI
Reactions of SX5E Returns to Changes in Environmental Uncertain Talking Measure (Hard Information versus Soft Information)

This table reports the results for regressions of SX5E returns on the changes in the environmental uncertain talking measure in answers. The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window. $\Delta EnvUncA\%$ is the first difference of the environmental uncertain talking measure in answers. It is calculated based on Equation (1). Words that we use to build up this measure are listed in the cluster *Risk* of Table VIII. The uncertain talking is a piece of soft information that signifies ECB's view on economic and financial environments as well as its monetary policy stance. Panel (a) compares the effect of uncertain talking on press conference days with and without a publication of ECB's staff macroeconomic projection, which is a piece of hard information on ECB's view on economic conditions. Panel (b) compares the effect of uncertain talking on press conference days with and without a monetary policy action, which is a piece of hard information on ECB's monetary policy stance. All the control variables listed on Table IV are included in the regressions. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) Stock Returns in Press Conferences with and without Projection Release

	Without Projection	With Projection
$\Delta EnvUncA\%$	0.202	0.436^{**}
	(0.130)	(0.191)
Controls	All	All
Observations	136	57
$R^2(\%)$	26.0	37.9

(b) Stock Returns in Press Conferences With and Without Policy Action

	Without Policy Action	With Policy Action
$\Delta EnvUncA\%$	0.303^{***}	0.567
	(0.104)	(0.712)
Controls	All	All
Observations	158	
R^2 (%)	17.6	71.1

Internet Appendix

Table of Contents

A Text Cleaning Procedure

We apply the following text cleaning procedure to prepare the textual data.

- Split each press conference transcript into three sub-components: the Introductory Statements, questions, and answers
- In the Introductory Statement, remove welcoming and concluding remarks
- Apply the Stanford's CoreNLP package (Manning et al., 2014) to detect named entities, as well as the parts-of-speech and base forms of each word
- Convert words into lower case
- Remove the uncertain word possible, if it appears in the expressions *as ... as possible*. This step deletes expressions such as *as soon as possible*, *as clearly as possible*, etc.
- Remove an uncertain word, if it is after a negation word such as *no* and *not*; This step deletes expressions such as *no ambiguity*, *not assume*, etc.
- Remove several types of named entities identified by CoreNLP, which include *date, duration, money, number, ordinal number, percent, time, URL address,* and *E-Mail address*. This step deletes expressions such as *the beginning of 2014*, *the next few months*, etc.
- Remove several types of parts-of-speech identified by CoreNLP, including

    1. determiners (e.g., *a*, *the*)
    2. predeterminers (e.g., *all*, *such*)
    3. Wh-determiners (e.g., *which*, *what*)
    4. the existential (*there*)
    5. personal pronouns (e.g., *you*, *we*)
    6. possessive pronouns (e.g., *yours*, *its*)
    7. Wh-pronouns (e.g., *what*, *who*)
    8. the possessive wh-pronoun (*whose*)

9. interjections (e.g., *yes*, *ok*)
10. prepositions (e.g., *of*, *in*)
11. particles (e.g., *up*, *out*)
12. Wh-adverbs (e.g., *when*, *where*)

- Remove stop words *to*, *do*, *have*, *be*, single letter words, and symbols

- Remove morphological affixes from words by performing a stemming procedure

B Annotated ECB Press Conference Transcripts

To see more examples of sentences including words with uncertain meanings please visit

https://zexisun.github.io/vauge_examples/index.html

On this page we annotated the uncertain words in each press conference transcript published by the ECB from 2000-2018. We further distinguished between weak modal words and moderate modal words from Loughran and McDonald (2011)'s word lists.

C A Non-Technical Introduction to the Word Vector

The vector representation of a word is introduced in Mikolov et al. (2013). The word vectors embed the semantic and syntactic patterns of the word and are generated from a neural network. In machine learning terminologies, these word vectors are obtained by training an unsupervised neural network model on textual data. From a computational economics point of view, training a neural network is similar to finding the solution of a nonlinear function starting from a random initial guess. We first represent each word as a randomized vector, and then repeatedly update these vector representations until they converge or until we reach the maximal number of iterations. In each iteration, we update the word vectors according to the *distributional hypothesis*, which states that words appearing in the same contexts tend to have similar meanings (Harris, 1954). In practice, it amounts to updating word vectors, so that the inner products of two nearby words' vectors are maximized compared with that of two distant words' vectors. After a sufficient number of iterations, each vector will capture the semantic and syntactic patterns of the corresponding word.

The key feature of word vector is that the cosine distance between any two word vectors quantifies the similarity of the meanings of the two corresponding words. Thus, if the cosine distance between the two word vectors is high, then the meanings of the two corresponding words should be similar. Therefore, grouping word vectors with low distances together is equivalent to grouping words with similar meanings together.

A direct approach to evaluating the performance of word vectors derived from the above method is to find verbal analogies. Denote $\mathbf{u}_X$ as the vector for the word X. If the relationship between the words A and B is similar to the relationship between the words P and Q, then the vector $\mathbf{u}_Q$ must have the largest cosine similarity to the vector $\mathbf{u}_A - \mathbf{u}_B + \mathbf{u}_P$ among all the word vectors. Indeed, using the word vectors, we can verify the equivalence of the following relationships (recall that all words are converted to lower case in the text cleaning procedure):

$$\textit{germany} \text{ to } \textit{bundesbank} \approx \textit{us} \text{ to } \textit{fed}$$

$$\textit{inflation} \text{ to } \textit{inflationary} \approx \textit{uncertainty} \text{ to } \textit{uncertain}$$

$$\textit{us} \text{ to } \textit{dollar} \approx \textit{china} \text{ to } \textit{renminbi}$$

D An Alternative Visualization of Word Vectors

In the main text, we employ the principal component analysis algorithm (PCA) to reduce the dimensionality of the word vectors and directly visualize these vectors in a two-dimensional diagram.

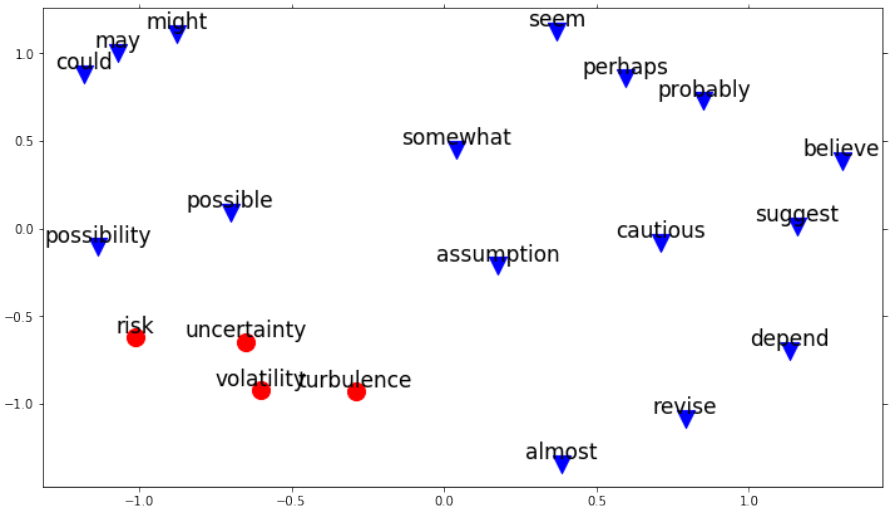
In this section, we take another approach that is introduced in Ulrich et al. (2019) to visualize the similarities (or equivalently the cosine distance) of the word vectors by employing the multidimensional scaling algorithm (MDS). To begin with, we calculate the cosine distance between any two of the first 20 uncertain words to obtain a 20-by-20 distance matrix. Each line is a distance vector that records the distances of one word vector to all the other word vectors. Then we apply the MDS algorithm to reduce the dimension of this distance matrix to two.

Intuitively speaking, the main difference between the PCA and MDS algorithms is that the PCA algorithm is applied to high-dimensional data and attempts to reduce dimensions while preserving the most variance. On the contrary, the MDS algorithm is applied to distances between objects and attempts to reduce dimensions while preserving the distance information.

The results are displayed in Figure 1. Words from clusters *Risk* and *Perhaps* are denoted as red dots and blue triangles, respectively. The distance between any two words within the same cluster is generally smaller than that from different clusters. This result further substantiates the intuition behind the K-Means algorithm we used.

Figure 1
Visualization of Word Vectors and Word Clusters

This figure visualizes the 20-by-20 distance matrix between the top 20 most-mentioned uncertain words. With Multidimensional Scaling (MDS) methods, the dimension of the distance matrix of word vectors is reduced to 2. Words from the cluster *Perhaps* are denoted as blue triangles, while words from the cluster *Risk* are denoted as red dots. For any two dots, the closer they are, the similar are their meanings.



E Daily Returns: Country- and Industry-Specific Stock Indices

We inspect the impacts of the uncertain talking measure on the daily stock returns on ten MSCI European country indices and display the results in Panel (a) of Table A1. The indices of nine out of ten countries react significantly and positively to the change in the uncertain talking in answers on press conference day. We also inspect the impacts of uncertain talking on the daily stock returns on 18 STOXX European industry indices and display the results in Panel (b) of Table A1. The positive reaction of stock return applies to 11 out of 18 industries, which includes those that are sensitive to macroeconomic and financial environments, such as banking, financial services, insurance, as well as the oil and gas industries.

Table A1
Reaction of Stock Returns to the Change in the Uncertain Talking Measure in Answers
(Daily Returns of Country- and Industry-Specific Stock Market Indices)

This table reports the results for regressions of daily returns of various stock market indices on the change in the uncertain talking measure in answers. The sample period is from Dec 2000 to Sep 2018. Panel (a) reports the results for the daily returns of ten MSCI Europe country indices. These countries are Austria, Belgium, Finland, France, Germany, Ireland, Italy, Netherland, Portugal and Spain. Panel (b) reports the results for eighteen STOXX Europe Total Market industry indices. The independent variable, $\Delta UncA\%$, is the change in the uncertain talking measure in presidents' answers to the ECB's press conference. All the control variables listed on Table IV of the main text are included in the regressions. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) Country Indices

	AUT	BEL	FIN	FRA	DEU	IRL	ITA	NLD	PRT	ESP
$\Delta UncA\%$	0.253	0.497**	0.425	0.455**	0.439*	0.470*	0.559***	0.337*	0.368**	0.364*
	(0.214)	(0.211)	(0.285)	(0.198)	(0.229)	(0.279)	(0.211)	(0.185)	(0.165)	(0.187)
Observations	194	194	194	194	194	194	194	194	194	194
$R^2(\%)$	11.7	10.9	16.1	15.3	15.3	23.2	18.9	13.6	11.7	14.9

(b) Industrial Indices

	Auto Parts	Banks	Basic Resource	Chemicals	Construction Materials	Financial Services	Food & Beverages	Health Care	Industry Goods
$\Delta UncA\%$	0.366	0.541**	0.404	0.511**	0.496***	0.382**	0.230	0.266	0.313
	(0.298)	(0.236)	(0.255)	(0.229)	(0.181)	(0.185)	(0.198)	(0.175)	(0.225)
Observations	194	194	194	194	194	194	194	194	194
$R^2(\%)$	10.2	16.4	14.1	13.9	14.5	15.2	9.7	17.3	14.4

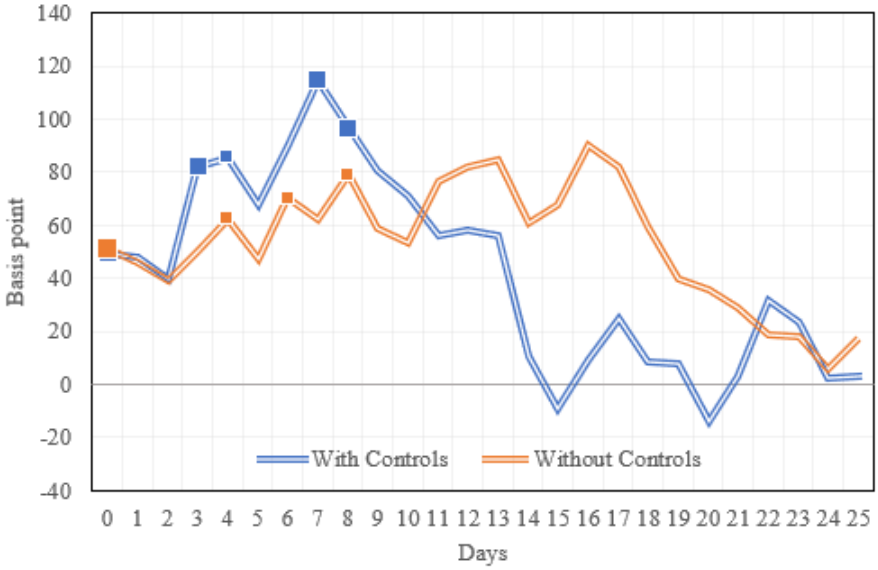
	Insurance	Media	Oil & Gas	Household Goods	Retail	Technology	Telecom	Travel & Leisure	Utilities
$\Delta UncA\%$	0.511**	0.332*	0.497**	0.401**	0.212	0.313	0.526***	0.508**	0.420**
	(0.235)	(0.187)	(0.237)	(0.190)	(0.174)	(0.252)	(0.193)	(0.206)	(0.173)
Observations	194	194	194	194	194	194	194	194	194
$R^2(\%)$	12.0	17.3	14.0	15.1	17.3	15.7	17.9	12.9	20.7

F Daily Returns: Persistence

To examine the persistence of the effect of uncertain talking, we regress the cumulative daily returns of the SX5E index. Figure 2 visualizes the estimated marginal effects of the uncertain talking measure in answers ($\Delta UncA\%$) on the cumulative returns of the SX5E on trading days after the press conference. Its significant impact and positive influence extends to the eighth and 15th day after the ECB's press conference, when control variables are included in the regression equation. This finding suggests that the effect of uncertain talking does not vanish immediately, but rather exists for an extended period after the press conference.

Figure 2
Reaction of SX5E Returns to the Change in the Uncertain Talking Measure in Answers (Daily Returns on Post Press Conference Days)

This figure displays the effects of the change in the uncertain talking measure in presidents' answers on the SX5E cumulative returns on the trading days after the press conference. The sample period is from Dec 2000 to Sep 2018. The cumulative return on the day N is the percentage change from the close price of the trading day before press conference day to the close price of the N-th trading day after press conference day. The blue line connects the estimated coefficients when no control variable is included in the regression. The orange line connects the estimation coefficients when all the control variables listed on Table IV are included in the regression. The large square marks the significance of the coefficient at 1% level, the middle is at 5% level, and the small is at 10% level.



G Daily Returns: All Trading Days

In the baseline analysis, we apply an event-study approach to concentrate on the reaction of the stock market returns on ECB's press conference days only. However, by adopting this approach, one cannot rule out the possibility that the results are driven by a systematic shift in returns on the ECB's press conference days, akin to what Lucca and Moench (2015) discovered for the excess returns before scheduled Federal Open Market Committee meetings. To check whether this phenomenon existed in our sample period and to further confirm the significant effect of uncertain talking, we regress daily returns of Eurozone stock market indices on all 4632 trading days from Dec 2000 to Sep 2018 and present the results in Table A2.

In the first specification, we add a dummy variable (PC), which equals one on press conference days and zero otherwise. Panel (a) shows that the effect of this dummy cannot significantly affect any of the stock returns listed in the table. Thus, we find no evidence suggesting a systematic shift in returns on press conference days. In the second specification, for each regressor considered in the baseline regression, we add an interaction term between this regressor and the dummy variable PC. As shown in Panels (b), the interaction term $PC \times \Delta UncA\%$ significantly and positively increases the returns of most stock market indices listed in the table.

Table A2
Reaction of Stock Returns to the Change in the Uncertain Talking Measure in Answers
(Daily Returns, All Trading Days)

This table extend the regression in the main text to all 4632 trading days from Dec 2000 to Sep 2018. The *PC* is a dummy that equals one when a ECB press conference takes place on that day, and zero otherwise. $PC \times \Delta UncA\%$ is the interaction term of the *PC* dummy and the change in the uncertain talking measure in presidents' answers on ECB press conference. Panel (a) presents the effect of the *PC* dummy on stock returns. Panel (b) shows the effect of the interaction term $PC \times \Delta UncA\%$ on stock returns, with all the control variables listed on Table IV of the main text included. $\Delta UncA\%$ and all the control variables are set to zero for non-press conference days. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	(1) STOXX 50 Daily	(2) STOXX Large	(3) STOXX Median	(4) STOXX Small	(5) MSCI	(6) MSCI Growth	(7) MSCI Value
Panel (a). PC Dummy							
PC	-0.036	-0.037	-0.006	0.036	-0.041	-0.043	-0.037
	(0.117)	(0.096)	(0.092)	(0.091)	(0.107)	(0.101)	(0.115)
Observations	4632	4632	4632	4632	4632	4632	4632
$R^2(\%)$	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Panel (b). Interaction							
$PC \times \Delta UncA\%$	0.519**	0.399**	0.337*	0.323*	0.463**	0.387**	0.525***
	(0.205)	(0.182)	(0.187)	(0.178)	(0.193)	(0.193)	(0.198)
Observations	4632	4632	4632	4632	4632	4632	4632
$R^2(\%)$	1.3	1.2	1.2	1.7	1.2	1.3	1.1

H Further Robustness Check: Outliers

To check the main findings depend only on a few points that have strongest influences, we remove observations whose impacts on the point estimate of the slope coefficient of $\Delta UncA\%$ are larger than a threshold. This is the standard practice to detect potential outliers and has been applied in other event-study analyses (e.g. Bernanke and Kuttner, 2005; Gorodnichenko and Weber, 2016). We use three thresholds to ensure the robustness of the results, which are 0.1, 0.15, and 0.3 times the standard error of the coefficient. Larger threshold means fewer observations to be removed from the sample. As displayed in Columns (3) to (8) of Table A3, the estimated response coefficient of SX5E returns is still positive and significant at the 1% level.

Table A3
Reaction of SX5E Returns to the Change in the Uncertain Talking Measure in Answers (Removing Outliers)

This table reports the regressions results of the SX5E returns on the change in the uncertain talking measure in answers after removing outliers. We also present the results when no outlier is deleted in Columns (1) and (2) as benchmarks. The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window. $\Delta UncA\%$ is the change in the uncertain talking measure in presidents' answers. Based on the DFBETA method, we identify outliers as those observations whose influences on the estimated slope coefficient of $\Delta UncA\%$ is larger than the thresholds. The threshold numbers are equal to 0.1, 0.15, or 0.3 times the standard error of the coefficient. We consider both the case when no control variable is included in the regression, and when all the control variables listed on Table IV of the main text are included. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	Original		Threshold = 0.1		Threshold = 0.15		Threshold = 0.3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta UncA\%$	0.212***	0.230***	0.204***	0.167***	0.194***	0.213***	0.173***	0.185***
	(0.066)	(0.078)	(0.046)	(0.062)	(0.046)	(0.067)	(0.047)	(0.066)
Controls	No	All	No	All	No	All	No	All
Observations	193	193	175	175	191	191	192	192
$R^2(\%)$	3.3	19.1	3.1	19.9	2.8	18.3	2.3	17.0

I Further Robustness Check: Alternative Uncertain Talking Measures and Control Variables

First, to exclude the possibility that the effect is significant only under a specific choice of the uncertain talking measure, we regress the SX5E return on variants of the standard uncertain talking measure and present the results in Panel (a) of Table A4.

As discussed in the definition of the uncertain talking measure, the uncertain talking measure in each part is defined as the number of uncertain words in that part scaled by the total number of words in the whole transcript. Here we consider two alternative choices: One is without the scaling parameter ($\Delta UncAScore$) and the other one is scaled by the total number of words in the corresponding part only ($\Delta UncAA\%$). Columns (1) and (2) of Panel (a) exhibit that these two measures still have significant and positive effects on the SX5E returns.

As discussed in the econometric strategy, the surprised component of the uncertain talking measure might not be its first difference. We consider two extremes here. If market participants have no expectations of the amount of uncertain talking, then the surprise component should equal the *level* of the uncertain talking measure. If otherwise, market participants formulate a sophisticated forecasting regression to predict the amount of uncertain talking, then the surprise component should equal the difference between the actual and predicted uncertain talking measures, or equivalently the *residual* of the forecasting regression. We construct a linear model to predict the uncertain talking measure based on variables that can be observed before the start of the press conference. Results from Columns (3)-(4) indicate that both the surprise components obtained under two extremes have a significant positive effect on stock returns.

Furthermore, in Panel (b) of Table A4, we demonstrate that the main finding is robust to the choices of alternative control variables. This includes changing the conventional policy control from the marginal refinancing operation rate to the deposit facility rate in Column (1), changing the unconventional policy control from the forward guidance dummy to the amount of assets purchased in the ECB's asset purchasing program in Column (2), and changing the readability control from the

total number of words to the total length of the document in Column (3), and finally changing the economic projection controls from nowcasts to forecasts in Column (4). In these specifications, the effect of $\Delta UncA\%$ remain almost unchanged.

Table A4
Reaction of SX5E Returns to the Change in the Uncertain Talking Measure in Answers (Alternative Uncertain Talking Measures and Control Variables)

This table reports the regression results of SX5E returns to alternative uncertain talking measures in Panel (a), and alternative control variables in Panel (b). The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window. In Panel (a), $\Delta UncAScore$ is the change in the number of uncertain words in presidents' answers. $\Delta UncAA\%$ is the change in the uncertain talking measure in presidents' answers, scaled by the total number of words in presidents' answers only. $UncA\%$ is the level of the uncertain talking measure in presidents' answers. $UncA\%Surprise$ is the residual of the uncertain talking measure based on a forecasting regression. Textual controls are adjusted accordingly to be consistent with new measures. In Panel (b): In Column (1) we replace ΔMRO with ΔDF, which is the change in the deposit facility rate. In Column (2) we replace FG with QE, which is the amount of assets the ECB purchases in its asset purchase program. In Column (3) we replace $\Delta TWords$ with $\Delta\%TLength$, which is the string length of the transcript. Column (4) displays the result when we replace GDP and HICP nowcast with one-year-ahead forecasts. Other control variables that are not explicitly shown in the table remain unchanged, as in Table IV of the main text. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) Alternative uncertain talking Measures

	(1)	(2)	(3)	(4)
$\Delta UncAScore$	0.009***			
	(0.003)			
$\Delta UncAA\%$		0.093**		
		(0.043)		
$UncA\%$			0.207*	
			(0.116)	
$UncA\%Surprise$				0.200*
				(0.118)
Observations	193	193	193	192
$R^2(\%)$	16.1	16.9	14.3	17.8

(b) Alternative Control Variables

	(1)	(2)	(3)	(4)
$\Delta UncA\%$	0.226***	0.246***	0.230***	0.222***
	(0.076)	(0.087)	(0.078)	(0.077)
ΔDF	-0.056			
	(0.477)			
QE		0.001		
		(0.002)		
$\Delta\%TLength$			-0.000	
			(0.00)	
$\Delta\%GDPForecast$				-0.004
				(0.065)
$\Delta\%HICPForecast$				0.046
				(0.120)
Observation	193	193	193	193
$R^2(\%)$	19.0	14.5	19.0	17.5

J Sub-Sample Analysis

We investigate the influence of uncertain talking in two sets of sub-samples and report the results in table A5. First, we split the samples into rate hike cycle and rate cut cycle, based on whether the most recent adjustment in an ECB's key policy rates, is negative, which signifies a tightening of the policy, or positive, which signifies an easing of the policy. The policy rate we choose is the deposit facility rate, lower of the three, because it continues to change after 2012, when the middle and higher rates are constrained at the zero lower bound. In addition, this is the most frequently adjusted policy rate of the ECB.

The results from this sub-sample regressions are displayed on the first two columns. The regression results show that the positive effect of uncertain talking is more prominent during the rate cut cycle.

We also present the results before and after 2009, when the global financial crisis occurred. The effect of uncertain talking is only significant during post-crisis period, and the scale of the coefficient is much larger than the one before the crisis. The results suggest that the unconventional positive response of stock return to uncertain talking grows after the crisis.

Table A5
Reaction of SX5E Returns to the Change in the uncertain talking Measure in Answers (Sub-sample Analysis)

This table reports the sub-sample results of regressing SX5E returns on the change of uncertain talking measure in presidents' answers. The sample period is from Dec 2000 to Sep 2018. The dependent variable is the SX5E return in the ECB's press conference time window, which is defined in Altavilla et al. (2019). $\Delta UncA\%$ is the change in the uncertain talking measure in presidents' answers in the ECB's press conference. A press conference is regarded as in the rate cut cycle if the most recent change in the deposit facility rate is negative, which signifies a tightening of the policy, otherwise in a rate hike cycle. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

	Policy Cycle		Crisis	
	Rate Hike	Rate Cut	Pre-2009	Post-2009
$\Delta UncA\%$	0.104	0.240***	0.076	0.374**
	(0.134)	(0.090)	(0.093)	(0.154)
Observations	45	148	103	90
$R^2(\%)$	53.0	23.8	17.3	34.1

K What Information Does the Environmental Uncertain Talking Reflect?

Based on the examples of sentences with words from the cluster *Risk* such as "risk", "uncertainty", "volatility", and "turbulence", it is reasonable to conjecture that, an increase in the environmental uncertain talking implies that the economic and financial environments become more uncertain and unfavorable in the view of the policymaker[1]. To confirm this conjecture, we examine whether the environmental uncertain talking are associated with the downward revisions of macroeconomic outlooks or heightened volatility of the financial markets, or both.

We first check whether the environmental uncertain talking reflects the revision of expectations of macroeconomic fundamentals. We look into the impact of the ECB's staff macroeconomic projection on changes in the environmental uncertain talking measure. The staff macroeconomic projection is published four times a year by the ECB, and it includes short- to long-term forecasts of real GDP growth, unemployment rates, inflation rates, and other aggregate variables. According to the description on the ECB website, it "contribute[s] to the ECB Governing Council's assessment of economic developments and risks to price stability" [2]. In other words, it is a crucial piece of information that the Governing Council will take into account when making the policy decision. Thus, the ECB president's uncertain talking might well be conditional on this piece of information.

We regress changes in the uncertain talking measures on the forecast of real GDP growth and HICP inflation rate on the current year and next year for the days when a staff projection is released, and report the results in Panel (a) of Table A6. A lowered GDP projection for next year significantly increases the environmental uncertain talking in presidents' answers, even when other factors, including lagged environmental uncertain talking, are controlled for. However, there are no significant influences on the attitudinal uncertain talking or the overall uncertain talking measure.

Next, we study whether uncertain talking reflects the variation of the financial markets. The uncertainty is represented by the Euro VSTOXX index, which is calculated based on the SX5E option

[1] We have removed those uncertain words following from a negation word in our text cleaning procedure.
[2] http://ecb.europa.eu/pub/projections/html/index.en.html

prices and has been used as a standard measure of the implied volatility in the stock markets of the Eurozone. According to the ECB's accounts of monetary policy meetings, financial markets developments is one of the major factors that the Governing Council reviews and takes into consideration when making the policy decision. Thus, there is a possibility that uncertainty in the financial markets might influence the ECB president's uncertain talking at the press conference. Besides, since the uncertain talking measure is built on a word list originally designed to gauge the uncertainty in financial reports, there is more reason to believe that the change in uncertain talking measures could be an indication of the ECB president's re-evaluation of financial market conditions.

The related results are presented in Panel (b) of Table A6. We find that the average percentage change in the VSTOXX index during the 20 trading days before the press conference indeed has a significant positive effect on the environmental uncertain talking, even when control variables are included in the regression. Similar to the results in Panel (a), there is no evidence to show that the change in the VSTOXX index can move the attitudinal uncertain talking or the overall uncertain talking measure.

To sum up, the analysis above reveals that the environmental uncertain talking indeed reflects the uncertain and unfavorable economic and financial environments faced by the ECB.

Table A6
Influence of External Environments on Different Uncertain Talking Measures in Answers

This table reports the influence of external environments the ECB faces on three different uncertain talking measures in answers. The sample period is from Dec 2000 to Sep 2018. The independent variables in the table are $\Delta EnvUncA\%$, $\Delta AttUncA\%$, and $\Delta UncA\%$. They are the changes in the environmental uncertain talking, attitudinal uncertain talking, and overall measure of uncertain talking in answers, respectively. Panel (a) shows the influence of macroeconomic environments, in which we use the ECB's staff projections on GDP growth and HICP inflation rates as proxies. $\Delta GDPf$, $\Delta HICPf$ are the first differences of one-year-ahead forecast of GDP and HICP. $\Delta GDPn$ and $\Delta HICPn$ are the first differences of nowcast of GDP and HICP for the current year. Panel (b) shows the influence of financial environments, in which we use $MPVSTX$, the average change in the VSTOXX index in the past 20 trading days as a proxy. The common control variables in these two panels are $L.\Delta EnvUncA\%$, $L.\Delta AttUncA\%$, and $L.\Delta UncA\%$, which are the first lags of the changes in the uncertain talking measures. Columns (1), (4), (7) in Panel (a) and Columns (1), (3) and (5) in Panel (b) present the results when no control variables is included. All other columns present the results when all the control variables listed in Table IV of the main text are included. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) uncertain talking Explained by ECB Economic Projection

	$\Delta EnvUncA\%$			$\Delta AttUncA\%$			$\Delta UncA\%$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta GDPf$	-0.102*	-0.123**	-0.133*	0.009	0.004	0.131	-0.054	-0.074	0.022
	(0.061)	(0.057)	(0.076)	(0.047)	(0.050)	(0.081)	(0.076)	(0.076)	(0.097)
$\Delta HICPf$	0.211**	0.174	0.133	-0.032	-0.123	-0.040	0.162	0.052	0.070
	(0.102)	(0.117)	(0.105)	(0.093)	(0.102)	(0.152)	(0.140)	(0.157)	(0.167)
$\Delta GDPn$		0.032	-0.026		-0.011	-0.052		0.013	-0.061
		(0.059)	(0.046)		(0.043)	(0.046)		(0.073)	(0.065)
$\Delta HICPn$		0.021	-0.062		0.148	0.086		0.153	0.009
		(0.118)	(0.087)		(0.091)	(0.087)		(0.168)	(0.126)
$L.\Delta EnvUncA\%$			-0.534***						
			(0.082)						
$L.\Delta AttUncA\%$						-0.377**			
						(0.141)			
$L.\Delta UncA\%$									-0.388***
									(0.130)
Controls	No	All	All	No	All	All	No	All	All
Observations	71	71	70	71	71	70	71	71	70
R^2 (%)	5.2	6.0	58.8	0.1	3.4	47.1	1.3	3.7	54.9

(b) uncertain talking Explained by Average VSTOXX Change in Past 20 Days

	$\Delta EnvUncA\%$		$\Delta AttUncA\%$		$\Delta UncA\%$	
	(1)	(2)	(3)	(4)	(5)	(6)
MPVSTX	0.046*	0.061***	-0.047	-0.002	0.007	0.063*
	(0.026)	(0.022)	(0.032)	(0.033)	(0.036)	(0.037)
$L.\Delta EnvUncA\%$		-0.437***				
		(0.057)				
$L.\Delta AttUncA\%$				-0.346***		
				(0.063)		
$L.\Delta UncA\%$						-0.332***
						(0.061)
Observations	194	193	194	193	194	193
Controls	No	All	No	All	No	All
R^2 (%)	1.6	40.5	1.3	28.1	0.0	32.0

L Reaction of the OIS Rates and German Government Bond Yields to the Overall Uncertain Talking Measure and Attitudinal Uncertain Talking Measure

We show in the main text that both the EONIA-based OIS rates and nominal German government bond yields decrease in response to an increase in the environmental uncertain talking measure. To confirm that such effects are only affected by the environmental uncertain talking, we check whether the overall uncertain talking measure or the attitudinal uncertain talking measure can significantly influence the OIS rates or German government bond yields.

The results are displayed in the last two panels of A7 and A8. The results from both tables confirm that there is no significant effect of either the overall uncertain talking or the attitudinal uncertain talking on the OIS rates or German government bond yields.

Table A7
Reaction of the OIS Rates to Changes in Uncertain Talking in Answers

This table reports the results for regressions of changes in the EONIA-based OIS (Overnight Indexed Swap) rates on the changes in environment, attitude, and overall uncertain talking in answers. The sample period is from Dec 2000 to Sep 2018. The independent variables in Panels (a), (b) and (c) are $\Delta EnvUncA\%$, $\Delta AttUncA\%$, and $\Delta UncA\%$. They are the changes in the environmental uncertain talking, attitudinal uncertain talking, and overall measure of uncertain talking in answers, respectively. Words in the *Risk* and attitudinal uncertain talkings are listed in the Panel (b) of Table VI of the main text. The dependent variables in Columns (1) to (10) of each panel are the changes in the OIS rates in the ECB's press conference time window, which is defined in Altavilla et al. (2019). The maturities of these OIS rates span from six months to 20 years. All the control variables listed on Table IV of the main text are included in the regressions. Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) Environment Uncertain Talking in Answers

	(1) OIS 6M	(2) OIS 1Y	(3) OIS 2Y	(4) OIS 3Y	(5) OIS 4Y	(6) OIS 5Y	(7) OIS 8Y	(8) OIS 10Y	(9) OIS 20Y
$\Delta EnvUncA\%$	-0.996	-1.708**	-1.766**	-1.947**	-1.489**	-1.490**	-0.986*	-1.037**	-1.024**
	(0.606)	(0.827)	(0.843)	(0.771)	(0.753)	(0.749)	(0.538)	(0.480)	(0.450)
Observations	193	193	193	193	193	193	193	193	167
R^2 (%)	10.3	11.6	13.4	14.3	13.1	12.9	14.8	15.3	22.4

(b) Attitude Uncertain Talking in Answers

	(1) OIS 6M	(2) OIS 1Y	(3) OIS 2Y	(4) OIS 3Y	(5) OIS 4Y	(6) OIS 5Y	(7) OIS 8Y	(8) OIS 10Y	(9) OIS 20Y
$\Delta AttUncA\%$	0.144	0.104	0.345	0.128	0.036	-0.361	-0.238	-0.308	-0.021
	(0.555)	(0.791)	(0.920)	(0.819)	(0.853)	(0.866)	(0.655)	(0.632)	(0.556)
Observations	193	193	193	193	193	193	193	193	167
R^2 (%)	9.1	9.9	12.0	12.2	11.8	11.7	13.8	14.1	20.6

(c) Overall Uncertain Talking in Answers

	(1) OIS 6M	(2) OIS 1Y	(3) OIS 2Y	(4) OIS 3Y	(5) OIS 4Y	(6) OIS 5Y	(7) OIS 8Y	(8) OIS 10Y	(9) OIS 20Y
$\Delta UncA\%$	-0.259	-0.567	-0.411	-0.602	-0.491	-0.670	-0.513	-0.589	-0.508
	(0.463)	(0.618)	(0.650)	(0.572)	(0.583)	(0.586)	(0.440)	(0.418)	(0.337)
Observations	193	193	193	193	193	193	193	193	167
R^2 (%)	9.3	10.3	12.1	12.6	12.1	12.2	14.4	14.9	21.4

Table A8
Reaction of the German Bond Yields to Changes in Uncertain Talking in Answers

This table reports the results for regressions of changes in the German government bond yields on the changes in the two dimensions as well as the overall measure of uncertain talking in answers. The sample period is from Dec 2000 to Sep 2018. The independent variables in Panels (a), (b) and (c) are $\Delta EnvUncA\%$, $\Delta AttUncA\%$, and $\Delta UncA\%$ respectively. They are the percentage changes in the environmental uncertain talking, attitudinal uncertain talking, and overall measure of uncertain talking in answers, respectively. Words in the environmental and attitudinal uncertain talking are listed in the Panel (b) of Table VI of the main text. The dependent variables in Columns (1) to (10) of each panel are the changes in the bond yields in the ECB's press conference time window, which is defined in Altavilla et al. (2019). The maturities of these bond yields rates span from six months to 30 years. All the control variables listed on Table IV of the main text are included in the regressions . Newey-West standard errors are shown in parentheses. ***Significant at 1%, **significant at 5%, *significant at 10%.

(a) Environment Uncertain Talking in Answers

	(1) DE6M	(2) DE1Y	(3) DE2Y	(4) DE3Y	(5) DE4Y	(6) DE5Y	(7) DE8Y	(8) DE10	(9) DE20	(10) DE30
$\Delta EnvUncA\%$	-1.111	-2.016**	-2.048**	-1.867**	-1.563**	-1.523*	-1.053*	-1.071*	-0.713	-0.974*
	(0.767)	(0.938)	(0.859)	(0.803)	(0.773)	(0.786)	(0.606)	(0.555)	(0.540)	(0.551)
Controls	All	All	All	All	All	All	All	All	All	All
Observations	134	163	193	193	193	193	193	193	193	193
R^2 (%)	9.7	11.4	13.0	12.2	13.1	12.8	13.5	13.2	13.9	12.5

(b) Attitude Uncertain Talking in Answers

	(1) DE6M	(2) DE1Y	(3) DE2Y	(4) DE3Y	(5) DE4Y	(6) DE5Y	(7) DE8Y	(8) DE10	(9) DE20	(10) DE30
$\Delta AttUncA\%$	-0.545	-0.328	-0.098	-0.030	-0.183	-0.565	-0.247	-0.279	-0.600	-0.210
	(0.661)	(0.716)	(0.959)	(0.947)	(0.923)	(0.942)	(0.793)	(0.774)	(0.656)	(0.675)
Controls	All	All	All	All	All	All	All	All	All	All
Observations	134	163	193	193	193	193	193	193	193	193
R^2(%)	8.8	8.7	11.2	10.5	11.8	11.8	12.6	12.1	13.9	11.5

(c) Overall Uncertain Talking in Answers

	(1) DE6M	(2) DE1Y	(3) DE2Y	(4) DE3Y	(5) DE4Y	(6) DE5Y	(7) DE8Y	(8) DE10	(9) DE20	(10) DE30
$\Delta UncA\%$	-0.750	-0.934	-0.790	-0.700	-0.675	-0.821	-0.579	-0.614	-0.641	-0.566
	(0.586)	(0.645)	(0.663)	(0.640)	(0.629)	(0.640)	(0.527)	(0.508)	(0.434)	(0.403)
Controls	All	All	All	All	All	All	All	All	All	All
Observations	134	163	193	193	193	193	193	193	193	193
R^2 (%)	9.6	10.0	11.8	11.0	12.3	12.3	13.2	12.9	14.5	12.2