# Words are not all created equal: a new measure of the ECB communication

Matthieu Picault Thomas Renault

January 9, 2017

Preliminary Version

**Abstract**

We develop a field-specific dictionary to measure the stance of the European Central Bank monetary policy (dovish, neutral, hawkish) and the state of the Eurozone economy (positive, neutral, negative) through the content of ECB press conferences. In contrast with traditional textual analysis, we propose a novel approach using term-weighting and contiguous sequence of words (n-grams) to better capture the subtlety of central bank communication. We find that quantifying ECB communication using our field-specific weighted lexicon do help predicting future ECB monetary decision and European stock market volatility. Our indicators significantly outperform a textual classification based on the Loughran-McDonald or Apel-Blix-Grimaldi dictionaries and a media-based measure of economic policy uncertainty.

## 1 Introduction

> *"What matters for transparency is therefore clarity as well as openness. For a new and supranational institution like the ECB, it is particularly important that it sends clear and coherent messages to the markets and the wider public."*
>
> Otmar Issing (Executive Board Member of the European Central Bank) - 1999

In recent years, central banks have moved toward a greater transparency (Geraats 2002) in terms of objectives, procedures, rationales, models, and data. Central bank now reveal more information to the public by (i) describing the strategy that guides policy decisions, (ii) explaining current policy decisions (iii) interpreting economic conditions and giving views on future economic outlook and (iv) making statements about future policy (Woodford 2005). In that regard, central bank communication has become a key instrument in the central bankers' toolbox, particularly during high uncertainty period or when interest rate reaches the zero lower bound (Filardo & Hofmann 2014). By managing expectations,

communication may improve the predictability of monetary policy and reduce volatility in financial markets (see Blinder et al. (2008) for a survey, and Sturm & De Haan (2011) among others).

Since its creation in 1998, the European Central Bank (ECB) have placed a strong focus on transparency, using various channel of communication to send "clear and coherent messages to the markets and the wider public". While interviews, speeches, press releases (and more recently webcasts and tweets) have developed overtime, the main channel of communication used by the ECB is the press conference held after the Government Council meeting. Every six weeks,[1] the President and the Vice President of the ECB explain their monetary policy decision and answer questions from journalist during a press conference. While every words pronounced by central banks officials are closely scrutinize by decision makers and market participants, converting qualitative content into quantitative indicators remains a challenge for researchers.

Building on previous works from textual analysis, we propose a new methodology to quantify ECB communication. Classifying manually all sentences in all ECB press conference between January 2006 and December 2014, we develop a field-specific lexicon to measure the stance of the ECB monetary policy (dovish, neutral, hawkish) and the Governing Council views on the Eurozone economic outlook (positive, neutral, negative). For each press conference, we compute a monetary policy indicator and an economic outlook indicator by analyzing words and group of words appearing in the introductory statement. Then, we assess if our indicators contain value-relevant information, not already measured by alternative quantification from the literature, about future monetary policy decisions or future European stock market return and volatility.

We find that quantifying ECB communication using a field-specific weighted lexicon do help predicting future ECB monetary decision and market volatility. A dovish (hawkish) textual content about monetary policy and a negative (positive) economic outlook both predict a dovish (restrictive) decision at the next ECB meeting. A dovish (hawkish) monetary policy and a positive (negative) economic outlook both predict a decrease (increase) in market volatility the day after ECB statement. We also provide several evidences showing that traditional approaches using equal-weighted single word lists such as the Loughran & McDonald (2011) financial dictionary (LM hereafter) and the Bennani & Neuenkirch (2017) central banking dictionary (BN hereafter) fail to capture the forward-looking content of the ECB introductory statement. We believe that researchers interested in quantifying central bank communication should pay a specific attention on the methodology used to derive quantitative indicator from qualitative textual content. To encourage further research in this area and enhance replicability, all data used in this paper are available online.[2]

Our paper is organized as follows. In Section 2, we review the literature on the quantification of central bank communication and its influence on the predictability of monetary

[1] ECB adjusted schedule of meetings from a four-weeks cycle prior to 2015 to a six-weeks cycle from now on.

[2] http://www.xxx-xxx.com

policy and financial markets. Section 3 details the data and describes how we create quantitative indicators of monetary policy and economic outlook from textual ECB statements. Section 4 reports our methodology and empirical findings about future monetary policy decision. Section 5 reports our methodology and empirical findings on market volatility and return. Section 6 presents robustness checks. Finally, Section 7 concludes.

## 2 Related Literature

This article relates to three strands of the literature on central bank communication: (i) its quantification through textual analysis, (ii) its influence on the predictability of monetary policy and (iii) its impact on asset prices and market volatility.

The first step of any quantitative analysis on central bank communication is to transform qualitative textual content into quantitative sentiment and/or topics data. For this purpose, two main methods have been used in the literature.[3]

Why not both?

First, central bank communication can be coded manually, following a narrative approach of monetary policy decisions proposed by Romer & Romer (1989). While researchers can convert communication into quantitative indicators on various topics, like for example the importance that policymakers assigned to reducing inflation relative to promoting real growth (Boschen & Mills 1995) or central bank views on exchange rate valuation (Dewachter et al. 2014), the most common classification consists of grading communication depending on monetary policy inclinations. Looking at the ECB, this approach was followed by Musard-Gies (2006) and Rosa & Verga (2007) who hand-coded each statement, according to the tone of the communication, into a discrete variable between +2 (very hawkish) and -2 (very dovish).[4] A similar methodology was implemented by Gerlach (2007) on ECB Monthly Bulletin.

As a next step toward a better understanding of communication, and in line with Kohn et al. (2004) findings on the importance of central bank communication on economic conditions and economic outlook, Berger et al. (2011) categorize the overall monetary policy stance on a scale from -3 (strong inclination to lower rates) to +3 (strong inclination to increase rates) using four subcategories: overall policy intention, price stability, real economy and monetary sector. Likewise, Conrad & Lamla (2010) classified each sentence of ECB statement into four categories (price developments, real economy, monetary aggregates, exchange rate) and three tendencies (positive, neutral, negative).

While manual classification is easy to implement, it presents several drawbacks. First, manual scoring is by definition subjective. For example, considering 62 ECB press conferences between 1999 and 2004, Carlo Rosa (rater #1) and Giovanni Verga (rater #2) (in

[3] Other methods used on the literature for topics detection include unsupervised topic classification (Latent Dirichlet Allocation in Jegadeesh & Wu (2015) or Latent Semantic Analysis in Boukus & Rosenberg (2006)). Machine learning methods for sentiment analysis have also been considered by Moniz & de Jong (2014))

[4] +1 (hawkish), 0 (neutral), -1 (dovish)

Rosa & Verga (2007)) disagree on 14 (22.58%) statements. Second, converting a document into a discrete class variable prevents from considering the smooth evolution of central bank communication.[5] Third, except when classified data are publicly available, results are not easily reproducible, limiting further research and comparability.

To solve - partly - those issues, another strand of the literature relies on dictionary-based and word-count approaches. The simplest example is provided by Jansen & De Haan (2007) who quantify communication regarding risks to price stability by simply counting the frequency of the word "*vigilance*" in ECB communications. A more standard approach consists in counting the number of positive and negative words in central bank communication using a pre-defined list of signed words from the Harvard IV-4 psychosociological dictionary or the LM financial dictionary. Using a bag-of-words approach as in Tetlock (2007), Jegadeesh & Wu (2015) convert FOMC meeting minutes into quantitative sentiment indicator, considering both Harvard IV and LM dictionaries. Schmeling & Wagner (2015) quantify ECB press conference by computing the ratio of negative words to total words using the LM financial dictionary. A similar methodology was used, among others, by Hansen et al. (2014), Cannon (2015) and Jansen et al. (2016).

However, given the specificity of central bank communication in terms of content, structure and topics discussed, quantifying communication using non field-specific lexicon may fail to capture all the dimensions and subtlety of central bank communication. Although the LM word lists have been increasingly popular in latest researches, content analysis can be further improved by constructing more authoritative and extensive field-specific dictionaries (Kearney & Liu 2014).

The second step of any quantitative analysis on central bank communication is to assess if value-relevant "soft information" can be extracted from textual content. To answer this question, the main methodology consists in adding a communication variable into traditional model (Taylor rule, asset pricing, asset volatility) to analyze empirically if communication improves our understanding (or forecast) of monetary policy or financial markets.

Regarding the predictability of future monetary policy decisions, several papers find that communication successfully conveys information not included in the available macroeconomic data. Rosa & Verga (2007) and Heinemann & Ullrich (2008) prove that including central bank communication improves the forecasts of ECB interest rate decisions from a Taylor (1993) rule model. These results hold even when forward looking macroeconomic variables and interbank interest rates are considered (Sturm & De Haan (2011)). Analyzing foreign exchange market, Conrad & Lamla (2010) find that the Euro currency appreciates against the US dollar in response to statements about increasing risks to price stability. Jansen & De Haan (2005) show that communication triggers an increase in the volatility while Dewachter et al. (2014) provide evidence of large jumps of the exchange rate for several hours after the release. Regarding equity market, Sadique et al. (2013) show that

[5] If the classification is done at the sentence level instead of the document level, nearly continuous variables can however be generated by aggregating/averaging across sentiment and topics.

Federal Reserve Beige Book tone affects stock market volatility and trading volume. At the intraday level Jegadeesh & Wu (2015) confirm that the tone of FOMC minutes helps predicting stock market volatility and returns. Schmeling & Wagner (2015) find that positive (negative) tone in ECB communication is associated with an increase (decrease) in stock prices and a lower (higher) volatility.

In this paper, we depart from the existing literature by proposing a novel methodology to quantify ECB communication. We provide several evidences showing that developing a field-specific lexicon significantly improves the predictability of future monetary policy. We also prove that disentangling content related to monetary policy from content related to the economic outlook improves forecast of both monetary policy and financial markets.

## 3 Quantifying ECB communication

To quantify ECB communication, we propose a novel methodology. We first manually classify all sentences in all ECB introductory statement on two categories (monetary policy and economic outlook) and three inclinations (positive, neutral and negative). Then, for each word (or group of words, n-grams hereafter) appearing in at least two ECB introductory statements, we compute the probability that this n-gram belongs to one of our two categories and three inclinations. Last, we compute the tone of each ECB statement by summing n-grams probabilities, using a weighted word count approach.

### 3.1 Field-specific lexicon generation

Several issues exist when popular lexicons used in the finance literature, such as the Harvard IV or the LM dictionary, are applied to quantify ECB communication. First, as Rosa & Verga (2007) and Berger et al. (2011) already pointed out, the ECB employs a very standardized form of communication, both in terms of structure and keywords used. Comparing words used in ECB communication over time, Amaya & Filbien (2015) document an increase in speeches similarity, consistent with a standardization of communication. In this regard, applying non-field specific lexicon may fail to capture all specificities of central bank communication. For example, in the LM dictionary,[6] the word "*downward*" is classified as negative while "*upward*" is not classified, whereas both words are perfect opposites[7] in the ECB's introductory statements. Second, considering single words (unigrams) rather than contiguous sequence of $n$ words (n-grams) might cause improper classification of tone. For example, "lower unemployment" (May. 2007) will be classified as negative using the LM dictionary due to the presence of the negative word "unemployment". Similarly, "risks to

[6] We used the latest update of their dictionary available on Bill McDonald's website http://www3.nd.edu/ mcdonald/Word_Lists.html

[7] Two examples : "the range for real GDP growth this year has been revised upwards" (Sept. 2010) and "the ranges for real GDP growth in 2011 and 2012 have been revised downwards." (Sept. 2011).

financial stability" (Sept. 2012) will be consider as positive using the LM dictionary due to the presence of the positive word "stability".

To address these limits, we generate a field-specific lexicon designed to quantify central bank communication. More precisely, we consider all sentences in all ECB introductory statements released between January 2006 and December 2014, covering 68 speeches from Jean-Claude Trichet and 38 speeches from Mario Draghi. For each 15,424 individual sentences, we follow a standard textual analysis methodology by (i) converting all words to lower case, removing numbers and punctuation (ii) using a Porter (1980) stemming algorithm to reduce inflected words to their word root (e.g, "increasing" to "increas", "unemployment" to "unemploy") (iii) removing a set of 32 stop words (e.g, a, the, an, of, to...). Then, following Kohn et al. (2004), we classify manually all 15,424 sentences pronounced during ECB introductory statement into seven categories: (1) monetary policy hawkish, (2) monetary policy neutral, (3) monetary policy dovish, (4) economic outlook positive, (5) economic outlook neutral, (6) economic outlook negative, (7) none.[8] The first three categories are grouped into a topic labeled *Monetary Policy* (MP) and refer to the ECB Governing council monetary policy decisions, without considering references to past decisions.[9] This topic also includes references to the short and medium term views of the Governing council on the expected path of monetary policies (see Appendix A.1 for selected sentences). The next three categories are grouped into a topic labeled *Economic Outlook* (EC hereafter, see Appendix A.2 for selected sentences) and focus on policy makers' description of the current economic situation and views on the future economic outlook. The last category (*None*, see Appendix A.3 for selected sentences) groups sentences not directly relevant to either monetary policy decisions or the Governing council economic outlook. This category also includes sentences presenting data that have already been released before ECB statement (HICP inflation, real GDP growth, monetary aggregates...) without any forward-looking statement nor additional information.

For each n-grams $n$ (from 1-gram to 10-grams) appearing at least twice in our sample, we count the frequency of occurrence of that n-gram in each of the seven categories defined previously, and we compute the probability that it belongs to category $c$ (MP and EC) with

[8] Each sentence is classified into the category that matters most, even though in few cases, a sentence may include information about monetary policy and about economic outlook. For example the sentence "Monetary developments therefore continue to require very careful monitoring, particularly against the background of improved economic conditions and continued strong property market developments in many parts of the euro area." (Jan. 2007) contains a "monetary policy hawkish" information justified by an "economic outlook positive" part. In this specific example, we consider that "what matters most" is the "monetary policy hawkish" tonality.

[9] For example, the sentence "The information that has become available since our last meeting has further underpinned the reasoning behind our decision to increase interest rates in [...]" appeared 6 times between November 2006 and August 2008. We consider this sentence as "not related to monetary policy nor economic outlook" as it do not convey new information to market participants

the inclination $i$ (dovish, neutral, hawkish for MP – positive, negative, neutral for EC).

$$P_n^{c,i} = \frac{number\ of\ occurrence\ _n^{c,i}}{total\ number\ of\ occurence\ _n} \tag{1}$$

Table 1 presents, for a selected sample of n-grams, the total number of occurrence (#) and the probabilities $P_n^{i,c}$ associated. For example, the bigram "consumption growth" appears 22 times in our sample: 20 times (91%) in sentences classified manually as "economic outlook positive" and 2 times (9%) in sentences classified as "economic outlook negative"[10].

**Table 1:** Term Frequency and Probabilities - Sample n-grams

| | | Monetary Policy | | | Economic Outlook | | | |
|---|---|---|---|---|---|---|---|---|
| n-grams | # | Dovi | Neut | Haw | Posi | Neut | Nega | None |
| act firm time manner ensur price stabil | 15 | 0.0 | 0.0 | 0.93 | 0.0 | 0.0 | 0.0 | 0.06 |
| remain present lower level extend period time | 11 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| purchas | 110 | 0.27 | 0.04 | 0.04 | 0.6 | 0.0 | 0.02 | 0.54 |
| cover bond purchas | 4 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| close readi consid all avail instrument | 3 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| decid reduc key ecb interest rate | 7 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| decid increas key ecb interest rate | 9 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| decid leav key ecb interest rate unchang | 22 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| uncertainti | 178 | 0.01 | 0.01 | 0.00 | 0.18 | 0.10 | 0.59 | 0.09 |
| uncertainti remain elev | 15 | 0.0 | 0.0 | 0.0 | 0.66 | 0.13 | 0.2 | 0.0 |
| uncertainti result turmoil financi | 6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| improv domest demand | 8 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| advers | 38 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.85 | 0.05 |
| develop | 799 | 0.06 | 0.15 | 0.11 | 0.14 | 0.06 | 0.18 | 0.31 |
| world economi | 39 | 0.0 | 0.0 | 0.0 | 0.46 | 0 | 0.46 | 0.08 |
| advers develop world economi | 8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| consumpt growth | 22 | 0.0 | 0.0 | 0.0 | 0.90 | 0.0 | 0.09 | 0.0 |
| lower consumpt growth | 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| been revis | 40 | 0.0 | 0.0 | 0.0 | 0.35 | 0.04 | 0.54 | 0.04 |
| been revis downward | 15 | 0.0 | 0.0 | 0.0 | 0.0 | 0.06 | 0.87 | 0.06 |
| been revis upward | 13 | 0.0 | 0.0 | 0.0 | 0.84 | 0.07 | 0.07 | 0.0 |
| been revis slightli upward | 3 | 0.0 | 0.0 | 0.0 | 0.75 | 0.0 | 0.0 | 0.25 |
| dampen | 102 | 0.06 | 0.00 | 0.0 | 0.07 | 0.03 | 0.65 | 0.16 |
| dampen underli growth momentum | 9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |

Notes: This table shows, for a list of selected n-grams, the total number of occurrence and the probabilities $P_n^{i,c}$ associated. For example, the word "uncertainty" was pronounced 178 times during ECB introductory statement between January 2006 and December 2014, of which 105 times (59%) in sentence associated with a "negative economic outlook". The 4-grams "uncertainti result turmoil financi" was pronounces 6 times and is always associated with a negative economic outlook.

We define our field-specific lexicon by considering only n-grams with a probability over

[10] "Consumption growth" was classified in sentence associated with a "negative economic outlook" during two speeches (Jan. 2010, Feb 2010): "In addition, low capacity utilisation rates are likely to dampen investment, and unemployment in the euro area is expected to increase somewhat further, thereby lowering consumption growth.". The trigram "lowering consumption growth" has a probability of 1 of being associated with a "negative economic outlook" in our lexicon, in such a way that all sentences are properly classified.

0.5 in one of our six class of interest (MP hawkish/neutral/dovish - EC positive/neutral/negative). With this exclusion, our final field-specific lexicon denoted $n$' is composed of 34,052 n-grams.[11]

### 3.2 Monetary Policy and Economic Outlook indicators

For a given speech $s$, we analyze all words and group of words pronounced by the ECB President and we consider a weighted word-count approach using our field-specific lexicon. More precisely, we define the probability for a speech $s$ of being classified in the category $c$ with the inclination $i$ as:

$$P_s^{c,i} = \frac{\sum_{n'=1}^{l} P_{n',s}^{c,i} * Occurrence_{n',s}}{\sum_{n'=1}^{l} P_{n',s}^{c} * Occurrence_{n',s}} \tag{2}$$

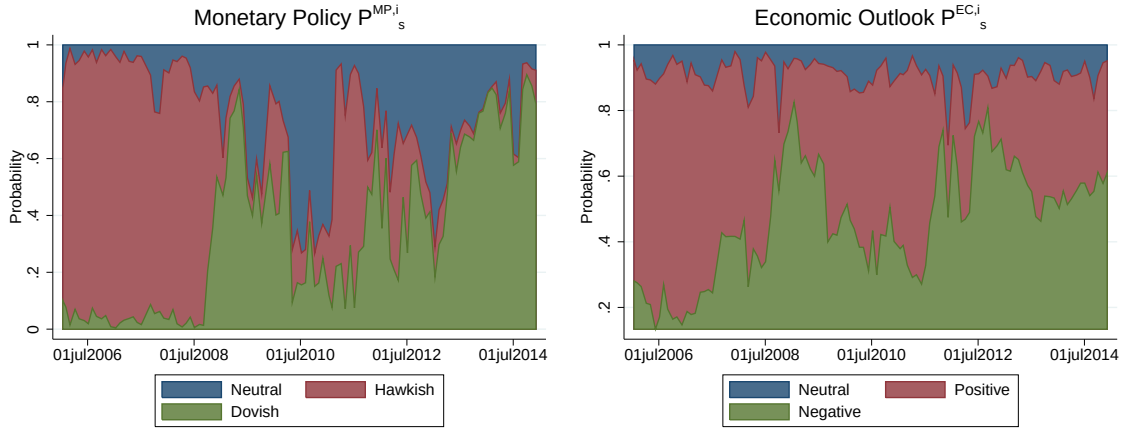
where l=34,052 (number of n-grams in our field-specific lexicon). For c = MP, i = *(hawkish, neutral, dovish)* and $\sum_{c=1}^{3} P_s^{MP,c} = 1$. For c = EC, i = *(positive, neutral, negative)* and $\sum_{c=1}^{3} P_s^{EC,c} = 1$. In order to improve the accuracy of our classification and to avoid multiple counting, we consider only the highest n-gram when multiple imbricated n-grams are found in a speech.[12]

Figure 1 shows, for all introductory statements between January 2006 and December 2014, the probabilities obtained from Equation 2 for our two categories of interest (Monetary Policy and Economic Outlook). Starting with the content related to monetary policy decisions and expected path of the monetary policy stance, our indicator is closely related to the evolution of the ECB monetary policy. First, from January 2006 to September 2008, a period during which ECB main refinancing rate increased from 2.25% to 4.25%, ECB communication about monetary policy was clearly hawkish. Then, starting in October 2008 and up to May 2010, the tone of the monetary policy became dovish. This period was associated with a large decrease in ECB key interest rate, from 4.25% to 1%. Communication then became neutral for few months, before a strong return of hawkish communication when ECB starts increasing its interest rate from March to September 2011. After that period,

[11] For example, we do not consider the bigram "world economy" as it appears 39 times, of which 18 times in EC positive (46.15%), 18 times in EC negative (46.15%), and 3 times in NONE (7.7%), in such a way that it does not convey clear information about the tonality of ECB communication by itself. However, the fourgram "adverse development world economy" is included in our lexicon as it appears 100% of the time in "EC negative" sentences.

[12] For example, as shown previously, both "consumption growth" and "lowering consumption growth" are part of our lexicon. However, in the sentence "thereby lowering consumption growth" (January and February 2010), we will consider only probabilities associated with the trigram "lowering consumption growth" ($P^{EC,nega} = 1$) without considering probabilities associated with the bigram "consumption growth".

**Figure 1:** Speech Probabilities for each category $c$ and inclination $i$




and up to the end of 2014, ECB communication became dovish, with both a decrease of the key interest rate and the implementation of various non-conventional monetary policy (LTRO, TLTRO, forward guidance, quantitative easing...).

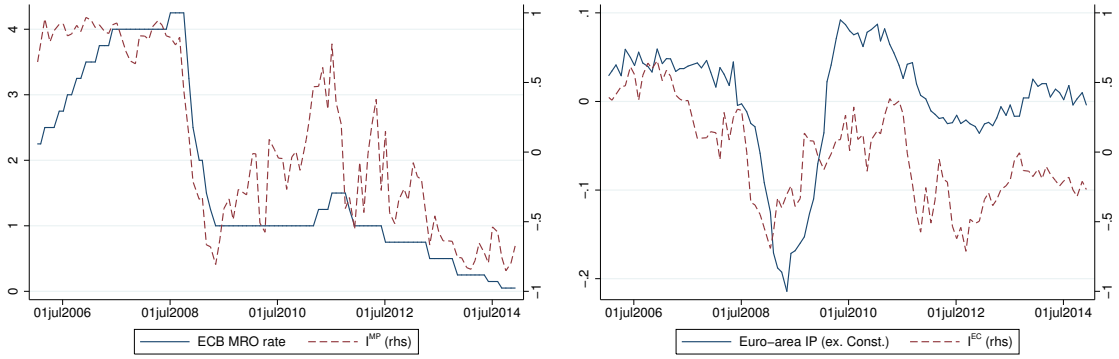
Regarding the economic outlook, our indicator capture both the subprime crisis and the eurozone crisis. Interestingly, the economic outlook starts deteriorating in September 2007, nearly one year before the Lehman Brothers bankruptcy, due to, amongst other things, "risk in financial markets on confidence" (Oct. 2007), "prolonged financial market volatility and re-pricing of risk on the real economy" (Nov. 2007) and "uncertainty about the potential impact on the real economy" (Dec. 2007). Regarding the eurozone crisis, the economic outlook became more and more negative starting in September 2011, characterized by a "moderation in the pace of global growth" (Sept. 2011), a "significant downward revision to forecasts" (Nov. 2011) and a "further intensification of the tensions in euro area financial markets" (Dec. 2011).

Finally, we aggregate the information content of the introductory statement into two indicators. To do so, we define two variables labeled $I^{MP}$ and $I^{EC}$ by computing, for each speech $s$, the difference between the hawkish (positive) probability $P_s^{MP,hawk.}$ ($P_s^{EC,posi.}$) and the dovish (negative) probability $P_s^{MP,dovi.}$ ($P_s^{EC,nega.}$).

$$s(I^{MP}, I^{EC}) = \begin{cases} I^{MP} = P_s^{MP,hawk.} - P_s^{MP,dovi.}, I^{MP} \in [-1,1] \\ I^{EC} = P_s^{EC,posi.} - P_s^{EC,nega.}, I^{EC} \in [-1,1] \end{cases} \tag{3}$$

Figure 2 displays the evolution of $I_t^{MP}$ compared to the ECB Main Refinancing Operation rate (MRR) and the evolution of $I_t^{EC}$ with the euro area industrial production (excluding construction).

**Figure 2:** ECB communication on Monetary Policy ($I_t^{MP}$) and Economic Outlook ($I_t^{EC}$)




### 3.3 Comparison with alternative measures of tone

We relate our new indicators to two alternative measures for the content of the ECB introductory statements. First, using the LM dictionary, we compute, for a given communication $s$, the tone $LM_s$ as the difference between the number of positive and negative words in the introductory statement divided by the total number of words identified. By construction, $LM_s \in [-1, +1]$ and is equal to 0 for a neutral speech. A positive (negative) value of $LM_i$ represents a statement with a relatively positive (negative) wording. Second, in line with Apel & Grimaldi (2012), Bennani & Neuenkirch (2017) divide a list of monetary policy relevant keywords into dovish or hawkish categories. With their classification adjusted for the content of introductory statements, we use a similar calculation than for $LM_s$ to assess the monetary inclination of the introductory statement. This measure is labeled $BN_s$. Figure 3 displays these two alternative measures with our new indicators. Table 2 provides the pairwise correlation coefficients and their significance level. On the overall sample, correlations are elevated (between 0.70 and 0.88) and significant at a 1% confidence level. The measure $LM_t$ captures more efficiently the information content related to the economic outlook while $BN_t$, following Bennani & Neuenkirch (2017) objective, successfully measures the monetary policy content of the introductory statements. However, starting mid-2011, the two alternative measures fail to consider negative communications following the sovereign debt crisis and the dovish tone in line with the non-standard policies implemented such as the 3 years LTRO, the forward guidance or the TLTRO.[13] Between June 2011 and December 2014, correlations between the different measures range from 0.35 to 0.66 but are still significant at a 1% or 5% confidence level.

[13] LTRO : Long term refinancing operations, TLTRO : Targeted LTRO.

**Figure 3:** Different measures for the content of ECB introductory statements

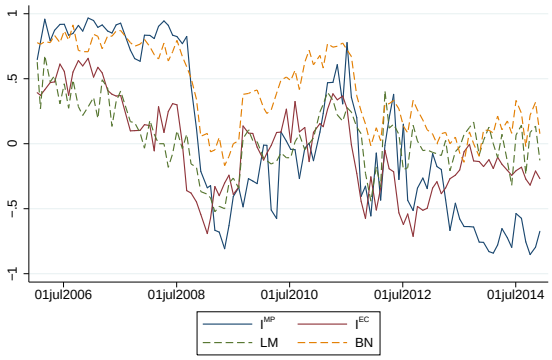



**Table 2:** Pearson correlations for the content of ECB introductory statements

| | Full Sample (106 obs.) | | | | June 2011 to Dec. 2014 (43 obs.) | | | |
|---|---|---|---|---|---|---|---|---|
| | $I_t^{EC}$ | $I_t^{MP}$ | $LM_t$ | $BN_s$ | $I_t^{EC}$ | $I_t^{MP}$ | $LM_t$ | $BN_s$ |
| $I_t^{EC}$ | 1 | | | | 1 | | | |
| $I_t^{MP}$ | 0.712*** | 1 | | | 0.174 | 1 | | |
| $LM_t$ | 0.702*** | 0.610*** | 1 | | 0.448*** | 0.351** | 1 | |
| $BN_s$ | 0.805*** | 0.884*** | 0.719*** | 1 | 0.380** | 0.661*** | 0.592*** | 1 |

# 4 Explaining and forecasting ECB monetary policy decisions

## 4.1 Methodology

To assess the relation between the information content of the introductory statements and monetary policy decisions, we test empirically the explanatory power of our two indicators ($I_t^{MP}$ and $I_t^{EC}$) compared to $LM_t$ and $BN_t$ on both contemporaneous and future monetary policy decisions. More precisely, we consider the following Equation:

$$ECB_t = \alpha + \beta_1 I_t^{MP} + \beta_2 I_t^{EC} + \rho ECB_{t-1} + \epsilon_t \tag{4}$$

where $ECB_t$ is the ECB monetary policy decision at time $t$, $I_t^{MP}$ and $I_t^{EC}$ are our indicators of communication, $\alpha$ is a constant and $\epsilon_t$ an error term. We include the lagged decision $ECB_{t-1}$ to control for the smoothing of monetary policy. In a forward looking approach, equation 4 can be rewritten as:

$$ECB_{t+m} = \alpha + \beta_1 I_t^{MP} + \beta_2 I_t^{EC} + \rho ECB_t + \epsilon_t \tag{5}$$

with $m = 1$ for the next Governing Council monetary policy decision and $m = 2$ the two-periods ahead decision. To control for available economic information at the time of the introductory statement, we consider a forward looking Taylor (1993) monetary policy rule (Orphanides (2001)) with contemporaneous and forward looking measures of inflation and output gap as in Jansen & De Haan (2009):

$$ECB_t = \alpha + \beta_1 I_t^{MP} + \beta_2 I_t^{EC} + \gamma_1(\pi_t - \pi^*) + \gamma_2(y_t - y^*) + \gamma_3\ \pi_t^e + \gamma_4\ y_t^e + \rho ECB_{t-1} + \epsilon_t \quad (6)$$

$$ECB_{t+m} = \alpha + \beta_1 I_t^{MP} + \beta_2 I_t^{EC} + \gamma_1(\pi_t - \pi^*) + \gamma_2(y_t - y^*) + \gamma_3\ \pi_t^e + \gamma_4\ y_t^e + \rho ECB_t + \epsilon_t \quad (7)$$

where $(\pi_t - \pi^*)$ is the inflation gap defined as the difference between the current level of inflation (Euro Area HICP) available at the time of the statement[14] and the ECB inflation target $\pi^* = 2\%$. The inflation expectations $\pi_t^e$ are proxied with the 12-month ahead inflation forecast from the ECB Quarterly Survey to Professional Forecasters (SPF). The output gap $(y_t - y^*)$ is measured by the difference between the Euro Area industrial production (excluding construction, see Gerlach (2007))[15] and the potential output $y^*$, the trend of the time serie obtained with an Hodrick-Prescott filter.[16] Following Sauer & Sturm (2007), the output gap expectations $y_t^e$ are derived from the European Commission Economic Sentiment Indicator (ESI) minus its long term average. The first difference of all macroeconomic variables are used in the estimation for stationarity. See Appendix A.5 for a graph presenting the macroeconomic variables and A.6 for descriptive statistics and correlations.

If the central bank communication does not provide any information additional to previously released macroeconomic data, then both our indicators $I_t^{MP}$ and $I_t^{EC}$ should not be significant in Equation 6 and 7. If central bank communication conveys relevant information, we expect a positive coefficient for $I_t^{MP}$: a more hawkish (dovish) communication should be associated with more hawkish (dovish) monetary policy. We also expect a positive coefficient for $I_t^{EC}$: an optimistic (pessimistic) economic outlook from the Governing council should be associated with a more hawkish (dovish) monetary policy.

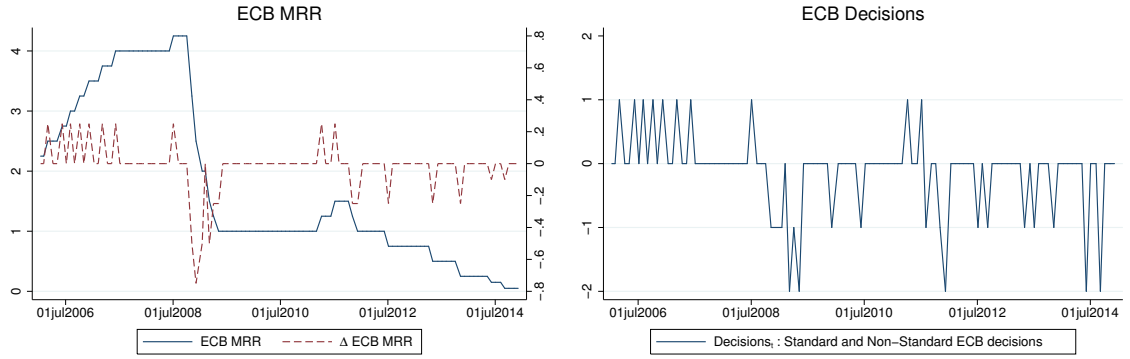
We consider two variables to measure the ECB monetary policy ($ECB_t$). First, we focus on interest rate decisions using the first difference of the Main Refinancing Operation interest rate (MRR hereafter, $\Delta MRR_t = (-0.75, -0.5, -0.25, -0.15, 0, +0.25)$). During our sample period, the Governing Council increased the MRR by 25 basis points on ten

[14] HICP inflation flash estimate is released at the end of the ongoing month but is subject to important revisions. The official HICP is available approximately 15 days after the end of the next month. To account for publication delay, we consider the official HICP data with a one month lag.

[15] Industrial production for a month $m$ is released around 13 days after the end of month $m + 1$. Again, to account for publication delay, we consider industrial production with a two month lag.

[16] The smoothing parameter $\lambda$ is set to 14,400.

**Figure 4:** Measures of the ECB monetary policy $ECB_t$




occasions and decreased it also on ten occasions (once by 75 basis points, once by 50 basis points, six times by 25 basis points and twice by 15 basis points). However, focusing on the MRR fails to consider non-standard policies implemented right after the beginning of the financial crisis. In the spirit of Jansen & De Haan (2009) and in order to account for the non-standard ECB policies, we create a variable $Decision_t$ (with $Decision_t = (-2, -1, 0, +1)$) taking the value of 0 when there is no change in the monetary policy stance, +1 for a hawkish monetary policy decision (an increase of the key interest rate by 25 basis points in our sample), -1 for a dovish monetary policy decision (either through a standard or a non-standard measure) and -2 for a very dovish decision with both a decrease of the key interest rate and a non-standard monetary measure. Appendix A.4 lists all the non-standard policies announced during ECB press conferences and considered in our sample period. Figure 4 presents both measures of ECB monetary policy ($\Delta MRR_t$ and $Decision_t$).

## 4.2 Empirical Findings

As our two monetary policy measures $ECB_t$ are discrete variables with 6 outcomes for $\Delta MRR_t$ and 4 outcomes for $Decision_t$, we use an ordered probit model to estimate the coefficient $\alpha, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3, \gamma_4$ and $\rho$ from Equations 4 to 7. We compare the performance of our indicators with two measures of tone using either the generic LM dictionary $LM_t$ or the central bank specific BN dictionary $BN_t$.

Table 3 and 4 summarize the results from a maximum likelihood estimation of Equations 4 and 6 (contemporaneous relationship). After controlling for both backward and forward looking macroeconomic variables, we find that $I_t^{EC}$ is significant at a 1% level for $\Delta MRR_t$ and $Decision_t$ while $I_t^{MP}$ is less significant (5%) when considering the change of the MRR. An hawkish/positive (dovish/negative) communication is associated with an increase (decrease) in ECB MRR and more hawkish (dovish) monetary policy decision (standard and

non-standard). As expected, we also find that the inflation gap ($\pi_t - \pi^*$) is significant at a 5% confidence level. However, we do not find that our indicators strongly improve the explanation of the current monetary policy compared to a sentiment indicator from the LM dictionary. Thus, when explaining current ECB monetary policy, using the LM dictionary seems to be sufficient to capture ECB sentiment. However, using forward looking macroeconomic variables, our indicators significantly improve the two alternative measures ($LM_t$ and $BN_t$) and previous results from Jansen & De Haan (2009).

Table 5 and 6 summarize the results from a maximum likelihood estimation of Equations 5 and 7 (future monetary policy decisions) for $m = 1$ and $m = 2$. We find that $I_t^{EC}$ is significant at a 1% confidence level while $I_t^{MP}$ does not convey relevant information to explain future policy decisions.[17] A positive (negative) economic outlook at time $t$ helps forecasting a hawkish (dovish) ECB policy at time t+1 and t+2. This result is consistent with Sturm & De Haan (2011) who empirically find that quantifying communication helps predicting the next policy decision of the ECB. Compared to the two alternative measures, our field-specific quantification of the content of ECB introductory statement significantly improves the predictability of future monetary policy decisions. Deriving tonality and topics using our field-specific lexicon provide a significantly better fit compared to a model where sentiment is computed using the $LM$ dictionary or the modified version of $BN$ words list.

# 5 Forecasting Stock Market

## 5.1 Methodology

In this section, we analyze stock market reactions to monetary policy statements. More precisely, we assess whether ECB communication explains the evolution of stock market return and volatility on statement days and/or predicts stock market return and volatility after ECB statement. For both contemporaneous relationship and forecast, we also analyze which components of ECB communication (monetary policy and/or economic outlook), if any, impact stock markets. For all regressions, we compare our results when ECB communication is quantified using the Loughran-McDonald and Bennani-Neuenkirch approaches.

We use daily closing value of the Eurostoxx50 (Eurostoxx) to compute stock market return and the "European VIX" (VSTOXX) for stock market volatility (Figure 5). We measure ECB decision surprise as the difference between the Bloomberg consensus prior to the decision and ECB rate announcement $MRR_t$.

$$Suprise_t = MRR_t - Consensus_t \tag{8}$$

To explain the link between monetary policy and stock market return on the day of the press conference (d=0) and on the day after the announcement (d=1), we consider the

[17] For $m = 1$ and when we consider $Decision_t$ as the dependent variable, $I_t^{MP}$ is significant at a 10% confidence level. In all other cases, $I_t^{MP}$ is not significant.

**Table 3:** Results from ordered probit models with the change of the ECB MRR ($\Delta MRR_t$): Estimation of Equations 4 and 6

| | Base Model | | | $I_t^c$ | | $LM_t$ | | $BN_t$ | |
|---|---|---|---|---|---|---|---|---|---|
| $\Delta MRR_{t-1}$ | 2.750*** | 2.226** | 0.583 | -0.420 | -2.240* | -1.069 | -1.702 | 0.521 | -0.708 |
| | (0.643) | (0.878) | (0.919) | (1.237) | (1.197) | (1.114) | (1.231) | (1.056) | (1.109) |
| $\Delta(y_t - y^*)$ | | 4.177 | -10.357 | 1.005 | -19.266 | -0.758 | -12.288 | 6.703 | -8.952 |
| | | (15.558) | (12.325) | (15.390) | (15.315) | (14.315) | (15.409) | (13.896) | (12.683) |
| $\Delta(\pi_t - \pi^*)$ | | 0.583 | 0.310 | 0.871 | 0.681 | 0.739 | 0.492 | 0.244 | 0.030 |
| | | (0.496) | (0.500) | (0.541) | (0.585) | (0.589) | (0.603) | (0.543) | (0.555) |
| $\Delta y_t^e$ | | | -0.022 | | -0.035 | | -0.034 | | -0.049 |
| | | | (0.084) | | (0.117) | | (0.113) | | (0.106) |
| $\Delta \pi_t^e$ | | | 4.292*** | | 5.225*** | | 3.019*** | | 3.947*** |
| | | | (0.949) | | (1.330) | | (1.115) | | (1.181) |
| $I_t^{MP}$ | | | | 1.143** | 1.675** | | | | |
| | | | | (0.496) | (0.767) | | | | |
| $I_t^{EC}$ | | | | 3.619*** | 4.525*** | | | | |
| | | | | (0.794) | (1.119) | | | | |
| $LM_t$ | | | | | | 5.704*** | 5.038*** | | |
| | | | | | | (0.773) | (0.809) | | |
| $BN_t$ | | | | | | | | 3.511*** | 3.680*** |
| | | | | | | | | (0.766) | (0.875) |
| Observations | 106 | 106 | 106 | 106 | 106 | 106 | 106 | 106 | 106 |
| $Pseudo - R^2$ | 0.0683 | 0.0777 | 0.193 | 0.321 | 0.429 | 0.346 | 0.389 | 0.237 | 0.316 |

Notes: The tables report the results from an ordered probit model estimated with maximum likelihood between January 2006 and December 2014. The dependent variable is the change of the ECB MRR. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

**Table 4:** Results from ordered probit models with ECB Policy Variable ($Decision_t$): Estimation of Equations 4 and 6

| | Base Model | | | $I_t^c$ | | $LM_t$ | | $BN_t$ | |
|---|---|---|---|---|---|---|---|---|---|
| $Decision_{t-1}$ | 0.245* | 0.166 | -0.131 | -0.634** | -0.903*** | -0.712*** | -0.815*** | -0.514** | -0.750*** |
| | (0.138) | (0.141) | (0.148) | (0.270) | (0.262) | (0.224) | (0.217) | (0.217) | (0.198) |
| $\Delta(y_t - y^*)$ | | 9.911 | -5.360 | 7.085 | -6.681 | 1.029 | -6.922 | 10.887 | -1.358 |
| | | (12.838) | (11.822) | (13.155) | (13.660) | (14.977) | (15.922) | (13.058) | (13.224) |
| $\Delta(\pi_t - \pi^*)$ | | 0.454 | -0.079 | -0.024 | -0.523 | -0.112 | -0.381 | -0.250 | -0.656 |
| | | (0.435) | (0.457) | (0.452) | (0.503) | (0.496) | (0.493) | (0.467) | (0.470) |
| $\Delta y_t^e$ | | | -0.035 | | -0.062 | | -0.022 | | -0.077 |
| | | | (0.063) | | (0.085) | | (0.069) | | (0.078) |
| $\Delta \pi_t^e$ | | | 3.387*** | | 2.966*** | | 1.851** | | 2.583*** |
| | | | (0.680) | | (0.834) | | (0.785) | | (0.801) |
| $I_t^{MP}$ | | | | 1.285*** | 1.266*** | | | | |
| | | | | (0.336) | (0.390) | | | | |
| $I_t^{EC}$ | | | | 2.129*** | 2.060*** | | | | |
| | | | | (0.615) | (0.626) | | | | |
| $LM_t$ | | | | | | 5.517*** | 5.074*** | | |
| | | | | | | (0.956) | (0.940) | | |
| $BN_t$ | | | | | | | | 3.425*** | 3.263*** |
| | | | | | | | | (0.561) | (0.588) |
| Observations | 106 | 106 | 106 | 106 | 106 | 106 | 106 | 106 | 106 |
| $Pseudo - R^2$ | 0.010 | 0.023 | 0.111 | 0.285 | 0.339 | 0.313 | 0.333 | 0.220 | 0.268 |

Notes: The tables report the results from an ordered probit model estimated with maximum likelihood between January 2006 and December 2014. The dependent variable represents ECB monetary policy decisions $Decision_t$. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

**Table 5:** Results from ordered probit models with the change of the ECB MRR ($\Delta MRR_{t+m}$): Estimation of Equations 5 and 7

| | m = 1 | | | | m = 2 | | | |
|---|---|---|---|---|---|---|---|---|
| $\Delta MRR_t$ | 0.931 | -1.257 | -0.083 | -0.420 | 1.841 | 0.016 | 1.139 | 1.505 |
| | (0.705) | (1.021) | (0.922) | (0.936) | (1.191) | (1.205) | (1.248) | (1.195) |
| $\Delta(y_t - y^*)$ | 21.160* | 15.239 | 20.199 | 22.721* | 13.802 | 7.860 | 12.525 | 14.050 |
| | (12.131) | (14.411) | (12.652) | (12.658) | (9.181) | (11.576) | (9.348) | (10.065) |
| $\Delta(\pi_t - \pi^*)$ | 0.007 | -0.092 | -0.033 | -0.272 | 0.805* | 0.892* | 0.813* | 0.687 |
| | (0.434) | (0.537) | (0.439) | (0.464) | (0.474) | (0.504) | (0.482) | (0.500) |
| $\Delta y_t^e$ | 0.083 | 0.138 | 0.092 | 0.091 | 0.091 | 0.121 | 0.097 | 0.095 |
| | (0.102) | (0.103) | (0.103) | (0.107) | (0.102) | (0.097) | (0.099) | (0.103) |
| $\Delta \pi_t^e$ | 2.321** | 2.576** | 1.776 | 1.429 | 0.287 | 0.276 | -0.101 | -0.300 |
| | (1.168) | (1.097) | (1.159) | (1.259) | (1.190) | (1.096) | (1.228) | (1.267) |
| $I_t^{MP}$ | | 0.004 | | | | -0.357 | | |
| | | (0.371) | | | | (0.348) | | |
| $I_t^{EC}$ | | 3.604*** | | | | 2.995*** | | |
| | | (0.722) | | | | (0.710) | | |
| $LM_t$ | | | 1.724** | | | | 1.194* | |
| | | | (0.744) | | | | (0.708) | |
| $BN_t$ | | | | 2.010*** | | | | 1.125*** |
| | | | | (0.566) | | | | (0.400) |
| Observations | 106 | 106 | 106 | 106 | 106 | 106 | 106 | 106 |
| $Pseudo-R^2$ | 0.139 | 0.326 | 0.175 | 0.209 | 0.0980 | 0.234 | 0.116 | 0.125 |

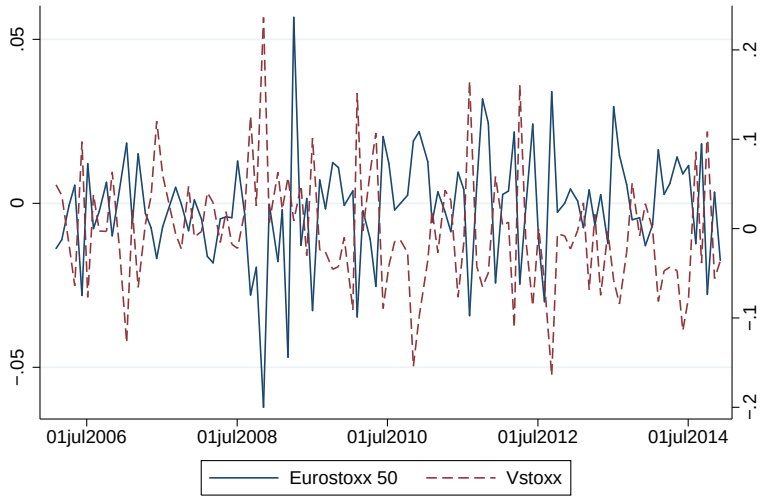
Notes: The tables report the results from an ordered probit model estimated with maximum likelihood between January 2006 and December 2014. The dependent variable is the one period ahead ($m = 1$) or two period ahead ($m = 2$) value of the change of the ECB MRR $\Delta MRR_{t+m}$. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

**Table 6:** Results from ordered probit models with ECB Policy Variable ($Decision_{t+m}$): Estimation of Equations 5 and 7

| | m = 1 | | | | m = 2 | | | |
|---|---|---|---|---|---|---|---|---|
| $Decision_t$ | -0.140 | -0.910*** | -0.653*** | -0.583*** | 0.339* | -0.026 | 0.163 | 0.165 |
| | (0.126) | (0.234) | (0.185) | (0.189) | (0.196) | (0.219) | (0.211) | (0.208) |
| $\Delta(y_t - y^*)$ | 15.972 | 10.023 | 14.341 | 17.240 | 13.651 | 9.457 | 12.135 | 13.611 |
| | (12.037) | (14.080) | (12.530) | (12.683) | (9.077) | (11.301) | (9.070) | (9.692) |
| $\Delta(\pi_t - \pi^*)$ | -0.132 | -0.539 | -0.271 | -0.559 | 0.658 | 0.461 | 0.605 | 0.479 |
| | (0.481) | (0.562) | (0.490) | (0.541) | (0.465) | (0.482) | (0.472) | (0.489) |
| $\Delta y_t^e$ | 0.123 | 0.080 | 0.084 | 0.017 | 0.031 | 0.023 | 0.020 | |
| | (0.089) | (0.087) | (0.087) | (0.093) | (0.088) | (0.084) | (0.087) | (0.088) |
| $\Delta \pi_t^e$ | 2.727*** | 2.607*** | 1.998** | 1.968** | 0.922 | 0.412 | 0.588 | 0.418 |
| | (0.830) | (0.921) | (0.801) | (0.974) | (0.855) | (0.835) | (0.887) | (0.895) |
| $I_t^{MP}$ | | 0.577* | | | | 0.256 | | |
| | | (0.313) | | | | (0.278) | | |
| $I_t^{EC}$ | | 2.592*** | | | | 1.596*** | | |
| | | (0.620) | | | | (0.590) | | |
| $LM_t$ | | | 2.711*** | | | | 0.997 | |
| | | | (0.701) | | | | (0.615) | |
| $BN_t$ | | | | 2.475*** | | | | 1.082*** |
| | | | | (0.506) | | | | (0.383) |
| Observations | 106 | 106 | 106 | 106 | 106 | 106 | 106 | 106 |
| $Pseudo - R^2$ | 0.0991 | 0.296 | 0.175 | 0.209 | 0.0855 | 0.172 | 0.0972 | 0.110 |

Notes: The tables report the results from an ordered probit model estimated with maximum likelihood between January 2006 and December 2014. The dependent variable is the one period ahead ($m = 1$) or two period ahead ($m = 2$) value of the ECB monetary policy decisions $Decision_{t+m}$. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

**Figure 5:** Variations of Stock Markets indicators




following model:

$$R_{t+d} = \alpha + \beta_1 * R_{t-d-1} + \beta_2 * Surprise_t + \beta_3 * I_t^{MP} + \beta_4 * I_t^{EC} + \epsilon_t \quad (9)$$

where $R_{t+d}$ represents the variation of the EuroStoxx50 on day $t+d$ relative to the announcement date $t$. On the press conference day ($d = 0$), and given previous results from Rosa (2011) on FOMC statement, we expect $\beta_2$ to be negative as a positive surprise about the main refinancing rate (i,e: a rate higher than expected) should lead to a decrease in stock market prices. In the same way, we expect $\beta_3$ to be negative if our monetary policy indicator incorporates information about future monetary policy stance not included in the surprise. We expect $\beta_4$ to be positive, as a good news about economic outlook should improve company's rationally discounted future cash flows.[18] On the day after the announcement ($d = 1$), we do not make any hypothesis about the significance or sign of the coefficients. If information about the current decision, the economic outlook and future monetary stance is correctly integrated into market closing price, we should not find any price predictability after the announcement.

But how can you disentangle good economic news from rate hikes?

To explain the link between monetary policy and stock market volatility, we consider a model similar to Equations 9 (replacing $R_{t+d}$ by $VOL_{t+d}$ and $R_{t+d-1}$ by $VOL_{t+d-1}$), and

[18] Bad news about economic outlook can also be good news for stock markets (Boyd et al. 2005) if investors anticipate a more dovish monetary policy in the future due to a worsening of the economic situation. However, we conjecture that the indirect effect - from bad economic outlook to dovish monetary policy - should already be captured by our MP indicator.

a model where we consider the absolute surprise (instead of the surprise) as in Rosa (2011):

$$VOL_{t+k,t+l} = \alpha + \beta_1 * VOL_{t+k-1} + \beta_2 * |Surprise_t| + \beta_3 * I_t^{MP} + \beta_4 * I_t^{EC} + \epsilon_t \quad (10)$$

where $VOL_{t+d}$ represents the variation of the VSTOXX index on day $t+d$ relative to the announcement date $t$. On the press conference day ($d = 0$), we expect $\beta_2$ to be positive as an unexpected decision should increase market volatility. We expect $\beta_3$ to be positive, as a more dovish monetary policy, especially during a period of high uncertainty as the 2008-2012 period, should decrease market volatility. We expect $\beta_4$ to be negative as a better economic outlook should reduce volatility. After the announcement ($d = 1$), we do not make any hypothesis about the significance nor the signs of the coefficients.

### 5.2 Empirical Findings

Table 7 presents our results for d=0 (contemporaneous relationship) for both the EuroStoxx and the VSTOXX. We also present results using the LM and BN dictionaries to quantify ECB communication.

We find that, similarly to Ehrmann & Fratzscher (2007) and Ranaldo & Rossi (2010), monetary policy communications significantly affects asset prices and volatility. Regarding the content of the introductory statement, our MP indicator is significant and of the expected sign at the 5% level for all models. When ECB statements about monetary policy are hawkish (dovish), stock market increases (decreases) and volatility decreases (increases) on announcement day. Our EC indicator is also significant at the 10% level for model [1] and [5]: a positive (negative) economic outlook is associated with higher (lower) stock market return and lower (higher) volatility. For both stock prices and volatility, we find that the approach we used to derive our MP and EC indicator significantly outperform sentiment based indicator derived by considering the Loughran-McDonald dictionary and the Apel-Blix dictionary. This finding reinforces our results from previous section on forecasting the monetary policy.

Then, we analyze if ECB statement at day $t$ helps predicting stock market at day $t+1$. Table 8 presents our results for d=1 for the EuroStoxx and the VSTOXX. We do not find any significant results when considering MP and EC to forecast stock return on the next trading day. Information seems to be instantaneously integrated into stock prices, in such a way that there is no predictability on the day following ECB announcement, consistent with the efficient market hypothesis. However, we find significant results regarding financial markets volatility. Economic Outlook and Monetary Policy indicators derived from ECB statement at date t helps predicting volatility at day t+1, respectively at the 5% level and at the 10% level.[19]

[19] Interestingly, while for d=0, monetary policy was significant at the 5% level and economic outlook only at the 10% level, the situation reverse for d=1. We conjecture that this result could be explained by

**Table 7:** Contemporaneous relationship regression results (d=0)

| | $R_t$ | | | $VOL_t$ | | | |
|---|---|---|---|---|---|---|---|
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] |
| Constant | -0.0006 | -0.0016 | -0.0001 | -0.0057 | -0.0028 | 0.0024 | 0.0003 |
| | (0.0014) | (0.0016) | (0.0025) | (0.0062) | (0.0063) | (0.0069) | (0.0103) |
| $Surprise_t$ | 0.0626 | 0.0562 | 0.0610 | -0.0793 | | | |
| | (0.0481) | (0.0524) | (0.0481) | (0.1377) | | | |
| $\|Surprise_t\|$ | | | | | -0.1966 | -0.2301** | -0.1929* |
| | | | | | (0.1207) | (0.1015) | (0.1105) |
| $R_{t-1}$ | -0.2206 | -0.1502 | -0.1696 | | | | |
| | (0.1359) | (0.1324) | (0.1370) | | | | |
| $VOL_{t-1}$ | | | | 0.0652 | 0.0866 | 0.1317 | 0.1246 |
| | | | | (0.0957) | (0.0906) | (0.0812) | (0.0840) |
| $I_t^{MP}$ | -0.0100*** | | | 0.0356** | 0.0342** | | |
| | (0.0034) | | | (0.0144) | (0.0150) | | |
| $I_t^{EC}$ | 0.0125* | | | -0.0413 | -0.0493* | | |
| | (0.0067) | | | (0.0290) | (0.0292) | | |
| $LM_t$ | | 0.0050 | | | | -0.0315 | |
| | | (0.0062) | | | | (0.0243) | |
| $BN_t$ | | | -0.0034 | | | | 0.0009 |
| | | | (0.0044) | | | | (0.0188) |
| Observations | 106 | 106 | 106 | 106 | 106 | 106 | 106 |
| $Adj. - R^2$ | 0.0726 | 0.0228 | 0.0216 | 0.0329 | 0.0519 | 0.0239 | 0.0116 |

Note: The table reports the results from a linear regression (d=0) of Equation 9 and Equation 10. The dependent variable on model [1], [2] and [3] is the percentage change of the Eurostoxx 50 on ECB statement days. The dependent variable on model [4], [5] and [6] is the percentage change of the VSTOXX on ECB statement day. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

However, this result holds only when ECB communication is measured using $I_t^{MP}$ and $I_t^{EC}$ indicators and disappears when communication is quantified through the LM and BN dictionaries. To identify the persistent impact of ECB communication on market volatility, the methodology used to derived quantitative forward-looking information from soft data is therefore of utmost importance. As "all words are not created equal", we provide empirical evidences showing that our weighted field-specific lexicon approach helps capturing all the subtlety of central bank communication and improves our understanding of the impact of communication on financial markets.

# 6 Robustness check

In this section, we first provide a robustness check showing the results of a real-time implementation of our methodology. Then, we present results comparing our indicator with two other measures of uncertainty used on the literature: the number of word related to "uncertainty" from the LM dictionary as in Jegadeesh & Wu (2015) and a media-based measure of economic policy uncertainty from Baker et al. (2015).

## 6.1 Real-time lexicon generation

In our methodology presented in Section 3, we classify all sentences in all ECB introductory statements from 2006 to 2014 in order to construct our field-specific lexicon. N-grams probabilities are computed on the full sample period, in such a way that $I_t^{MP}$ and $I_t^{EC}$ indicators are in reality ex-post measures. Sentences classified after a period $t$ may impact n-grams probabilities in $t$. To check the robustness of our indicators, we simulate a real-time implementation of our methodology, where only sentences classified in $t$ are used to compute $P_n^{i,c}$ from Equation 1. This approach is equivalent to a situation where a human coder analyzes and classifies each sentence of a speech when it is pronounced before updating n-grams probabilities and computing $P_n^{i,c}$. We denote those two real-time (unrevised) indicators $RT_t^{MP}$ and $RT_t^{EC}$.

Figure 6 presents together the full lexicon indicators $I_s^c$ and the real-time lexicon indicators $RT_s^c$. To confirm previous findings on the predictability of monetary policy, we replace our initial measures by their real-time equivalent in the empirical estimations from Section 4. Appendix A.7 and A.8 presents a summary of the results. We find that the economic content of the introductory statements remains significant at a 1% confidence level. A negative (positive) real-time economic outlook predicts a dovish (hawkish) monetary policy decision at the next ECB meeting.

However, the monetary policy indicator $RT_s^{MP}$ is no longer significant. This finding is consistent with the fact that our real-time monetary indicator underestimates the dovish

differences in the speed at which market participants' process "soft information" included in ECB statement, focusing first on monetary policy and monetary stance, and then more slowly incorporating information related to the economic outlook. We encourage further research in this area.

**Table 8:** Next day regression results (d=1)

| | $R_{t+1}$ | | | $VOL_{t+1}$ | | | |
|---|---|---|---|---|---|---|---|
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] |
| Constant | | 0.0020 | 0.0020 | -0.0071 | -0.0065 | -0.0056 | -0.0045 |
| | (0.0013) | (0.0014) | (0.0025) | (0.0044) | (0.0046) | (0.0049) | (0.0081) |
| $Surprise_t$ | -0.0192 | -0.0190** | -0.0200* | 0.0366 | | | |
| | (0.0118) | (0.0084) | (0.0109) | (0.0420) | | | |
| $\|Surprise_t\|$ | | | | | -0.0595 | -0.0088 | -0.0286 |
| | | | | | (0.0554) | (0.0435) | (0.0436) |
| $R_t$ | 0.1456 | 0.1718* | 0.1659 | | | | |
| | (0.1117) | (0.1007) | (0.1076) | | | | |
| $VOL_t$ | | | | 0.0688 | 0.0595 | 0.0944 | 0.0890 |
| | | | | (0.0725) | (0.0717) | (0.0676) | (0.0701) |
| $I_t^{MP}$ | -0.0036 | | | 0.0172* | 0.0174* | | |
| | (0.0027) | | | (0.0095) | (0.0094) | | |
| $I_t^{EC}$ | 0.0062 | | | -0.0349** | -0.0378** | | |
| | (0.0051) | | | (0.0172) | (0.0184) | | |
| $LM_t$ | | -0.0046 | | | | 0.0154 | |
| | | (0.0062) | | | | (0.0208) | |
| $BN_t$ | | | -0.0004 | | | | -0.0010 |
| | | | (0.0043) | | | | (0.0152) |
| Observations | 106 | 106 | 106 | 106 | 106 | 106 | 106 |
| $Adj. - R^2$ | 0.0201 | 0.0218 | 0.0153 | 0.0111 | 0.0133 | -0.0064 | -0.0119 |

Note: The table reports the results from a linear regression (d=1) of Equation 9 and Equation 10 The dependent variable on model [1], [2] and [3] is the percentage change of the Eurostoxx 50 on the day after ECB statement. The dependent variable on model [4], [5] and [6] is the percentage change of the VSTOXX on the day after ECB statement. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

tonality of ECB communication after Lehman Brothers bankruptcy (from September 2008 to September 2009). As in any supervised learning approach, a sufficient number of observations (classified sentences) is necessary to derive n-grams weights and to capture correctly, in real-time, the tonality of ECB communication.

**Figure 6:** Real-time introductory statement indicators


## 6.2 Alternative measure of uncertainty

We also consider two alternatives text-based measures of uncertainty in order to confirm (invalidate) our results on explaining (forecasting) market volatility. More precisely, we compute a measure of uncertainty by counting, in each ECB statement, the number of word included in the "uncertain words list" from the Loughran-McDonald dictionary (as in Jegadeesh & Wu (2015)). We denote this indicator $ULM_t$. We also consider the European media-based measure of economic policy uncertainty from Baker et al. (2015).[20] We denote this indicator $UBB_t$.

Table A.9 presents results from Equation 10 ($d = 0$) where we compare our indicators $I_t^{MP}$ and $I_t^{EC}$ to $ULM_t$ and $UBB_t$. The results confirm that using our field-specific lexicon approach provides a better proxy of market uncertainty around ECB communication compared to other text-based indicators used on the literature. $ULM_t$ and $UBB_t$ do not successfully explain the evolution of market volatility on ECB announcement days. On unreported test, we also find that our approach gives better results after the announcement (forecasting volatility).

[20] http://www.policyuncertainty.com/media/Europe_Policy_Uncertainty_Data.xlsx - The European Uncertainty index is computed by counting the frequency of uncertainty-related words in news reports from Le Monde and Le Figaro for France, Handelsblatt and Frankfurter Allgemeine Zeitung for Germany, Corriere Della Sera and La Repubblica for Italy, El Mundo and El Pais for Spain, and The Times of London and Financial Times for the United Kingdom.

# 7 Conclusion

Central bank communication has become a key instrument in the central bankers' toolbox. However, deriving quantitative indicators from soft textual data remains a challenging issues for both practitioners and academics. In this paper, we propose a novel approach using term-weighting and n-grams to better capture the subtlety of central bank communication We develop a publicly available field-specific lexicon to measure the stance of the monetary policy (dovish, neutral, hawkish) and the Governing Council view on the Eurozone economy (positive, neutral, negative). Computing two indicators at a speech level, we construct a continuous time series quantifying the tone of the ECB communications between 2006 and 2014.

We find that the content of the introductory statements help predicting future ECB standard and non-standard monetary decisions even after controlling for both backward and forward looking macroeconomic variables. A dovish (hawkish) textual content about monetary policy and a negative (positive) economic outlook both predict a dovish (restrictive) decision at the next ECB meeting. Quantifying ECB communication also helps forecasting market volatility. A hawkish (dovish) textual content about monetary policy and a negative (positive) economic outlook predict an increase (decrease) in market volatility the day after ECB statement. Our indicators significantly outperform a textual classification based on the Loughran-McDonald financial dictionary, on the Bennani-Neuenkirch dictionary, and a media-based measure of economic policy uncertainty.

Our results also shed light on the fact that researchers should be very cautious when relying on existing words list to quantify central bank structured communication. As words are not all created equal, we provide evidences that developing a field-specific weighted lexicon helps capturing the forward-looking information contained in central banks communication.

## References


Amaya, D. & Filbien, J.-Y. (2015), 'The similarity of ECB's communication', *Finance Research Letters* **13**, 234–242.

Apel, M. & Grimaldi, M. (2012), 'The information content of central bank minutes', *Riksbank Research Paper Series* (92).

Baker, S. R., Bloom, N. & Davis, S. J. (2015), Measuring economic policy uncertainty, Technical report, National Bureau of Economic Research.

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

Berger, H., de Haan, J. & Sturm, J.-E. (2011), 'Does money matter in the ecb strategy? new evidence based on ecb communication', *International Journal of Finance & Economics* **16**(1), 16–31.

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

Boschen, J. F. & Mills, L. O. (1995), 'The relation between narrative and money market indicators of monetary policy', *Economic inquiry* **33**(1), 24–44.

Boukus, E. & Rosenberg, J. V. (2006), 'The information content of FOMC minutes', *Working Paper* .

Boyd, J. H., Hu, J. & Jagannathan, R. (2005), 'The stock market's reaction to unemployment news: Why bad news is usually good for stocks', *The Journal of Finance* **60**(2), 649–672.

Cannon, S. (2015), 'Sentiment of the FOMC: Unscripted', *Economic Review-Federal Reserve Bank of Kansas City* p. 5.

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

Dewachter, H., Erdemlioglu, D., Gnabo, J.-Y. & Lecourt, C. (2014), 'The intra-day impact of communication on euro-dollar volatility and jumps', *Journal of International Money and Finance* **43**, 131–154.

Ehrmann, M. & Fratzscher, M. (2007), 'Communication by central bank committee members: Different strategies, same effectiveness?', *Journal of Money, Credit and Banking* **39**(2-3), 509–541.

Filardo, A. J. & Hofmann, B. (2014), 'Forward guidance at the zero lower bound', *BIS Quarterly Review March* .

Geraats, P. M. (2002), 'Central bank transparency', *The Economic Journal* **112**(483), F532–F565.

Gerlach, S. (2007), 'Interest rate setting by the ECB, 1999-2006: Words and deeds', *International Journal of Central Banking* **3**(3), 1–46.

Hansen, S., McMahon, M. & Prat, A. (2014), 'Transparency and deliberation within the fomc: a computational linguistics approach', *Center for Economic Performance* .

Heinemann, F. & Ullrich, K. (2008), 'Does it pay to watch central bankers' lips? the information content of ecb wording', *Swiss Journal of Economics* pp. 05–070.

Jansen, D.-J. & De Haan, J. (2005), 'Talking heads: the effects of ECB statements on the euro–dollar exchange rate', *Journal of International Money and Finance* **24**(2), 343 – 361. Exchange Rate EconomicsExchange Rate Economics.

Jansen, D.-J. & De Haan, J. (2007), 'The importance of being vigilant: Has ECB communication influenced euro area inflation expectations?'.

Jansen, D.-J. & De Haan, J. (2009), 'Has ECB communication been helpful in predicting interest rate decisions? an evaluation of the early years of the economic and monetary union', *Applied Economics* **41**(16), 1995–2003.

Jansen, D.-J., Moessner, R. et al. (2016), 'Communicating dissent on monetary policy: Evidence from central bank minutes', *DNB Working Paper Series* (512).

Jegadeesh, N. & Wu, D. (2015), 'Deciphering fedspeak: The information content of FOMC meetings', *Working Paper* .

Kearney, C. & Liu, S. (2014), 'Textual sentiment in finance: A survey of methods and models', *International Review of Financial Analysis* **33**, 171–185.

Kohn, D. L., Sack, B. P. et al. (2004), 'Central bank talk: does it matter and why?', *In: Bank of Canada (Ed.), Macroeconomics, Monetary Policy and Financial Stability* pp. 175–206.

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

Moniz, A. & de Jong, F. (2014), Predicting the impact of central bank communications on financial market investors' interest rate expectations, *in* 'European Semantic Web Conference', Springer, pp. 144–155.

Musard-Gies, M. (2006), 'Do ECBs statements steer short-term and long-term interest rates in the euro zone?', *The Manchester School* **74**.

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

Porter, M. F. (1980), 'An algorithm for suffix stripping', *Program* **14**(3), 130–137.

Ranaldo, A. & Rossi, E. (2010), 'The reaction of asset markets to swiss national bank communication', *Journal of International Money and Finance* **29**(3), 486 – 503.

Romer, C. D. & Romer, D. H. (1989), Does monetary policy matter? a new test in the spirit of friedman and schwartz, *in* 'NBER Macroeconomics Annual 1989, Volume 4', MIT Press, pp. 121–184.

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

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

Sadique, S., In, F., Veeraraghavan, M. & Wachtel, P. (2013), 'Soft information and economic activity: Evidence from the beige book', *Journal of Macroeconomics* **37**, 81–92.

Sauer, S. & Sturm, J.-E. (2007), 'Using taylor rules to understand european central bank monetary policy', *German Economic Review* **8**(3), 375–398.

Schmeling, M. & Wagner, C. (2015), 'Does central bank tone move asset prices?', *Working Paper* .

Sturm, J.-E. & De Haan, J. (2011), 'Does central bank communication really lead to better forecasts of policy decisions? new evidence based on a taylor rule model for the ECB', *Review of World Economics* **147**(1), 41–58.

Taylor, J. B. (1993), 'Discretion versus policy rules in practice', *Carnegie-Rochester Conference Series on Public Policy* **39**(0), 195 – 214.

Tetlock, P. C. (2007), 'Giving content to investor sentiment: The role of media in the stock market', *The Journal of Finance* **62**(3), 1139–1168.

Woodford, M. (2005), 'Central bank communication and policy effectiveness', *NBER Working Paper Series* p. 11898.

# A Appendix

## A.1 ECB Statement sentences classification : Topic Monetary Policy

| | |
|---|---|
| **Accomodative** | |
| 04/12/2008 | On the basis of its regular economic and monetary analyses, the Governing Council decided to reduce the key ECB interest rates by a further 75 basis points. |
| 02/10/2013 | The Governing Council confirms that it expects the key ECB interest rates to remain at present or lower levels for an extended period of time. |
| 13/01/2011 | Accordingly, the Governing Council will continue to monitor all developments over the period ahead very closely. |
| 04/09/2014 | The newly decided measures, together with the targeted longer term refinancing operations which will be conducted in two weeks, will have a sizeable impact on our balance sheet. |
| **Neutral** | |
| 02/11/2006 | On the basis of our regular economic and monetary analyses, we decided at today s meeting to leave the key ECB interest rates unchanged. |
| 05/11/2009 | The current rates remain appropriate. |
| 04/12/2014 | In this context, early next year the Governing Council will reassess the monetary stimulus achieved, the expansion of the balance sheet and the outlook for price developments. |
| **Restrictive** | |
| 05/10/2006 | At today s meeting, we decided to increase the key ECB interest rates by 25 basis points. |
| 06/07/2006 | Therefore, if our assumptions and baseline scenario are confirmed, a progressive withdrawal of monetary accommodation remains warranted. |
| 06/09/2007 | Accordingly, the Governing Council will monitor very closely all developments. |
| 14/01/2010 | The Governing Council will also continue to implement the gradual phasing out of the extraordinary liquidity measures that are not needed to the same extent as in the past. |

## A.2 ECB Statement sentences classification : Topic Economic Outlook

| | |
|---|---|
| **Positive** | |
| 07/12/2006 | Domestic demand in the euro area is expected to maintain its relatively strong momentum. |
| 05/03/2009 | Over the course of 2010, the economy is expected to gradually recover. |
| 07/07/2011 | Euro area exports should continue to be supported by the ongoing expansion in the world economy. |
| 01/08/2013 | Furthermore, the overall improvements in financial markets seen since last summer appear to be gradually working their way through to the real economy, as should the progress made in fiscal consolidation. |
| **Neutral** | |
| 03/08/2006 | Turning to price developments, according to Eurostat s flash estimate, annual HICP inflation was 2.5 in July 2006, unchanged from June and May. |
| 05/07/2007 | The risks surrounding this favourable outlook for economic growth are broadly balanced over the shorter term. |
| 03/12/2009 | The Governing Council continues to view the risks to this outlook as broadly balanced. |
| **Negative** | |
| 06/12/2007 | However, the reappraisal of risk in financial markets is still evolving and is accompanied by continued uncertainty about the potential impact on the real economy. |
| 06/11/2008 | To sum up, the intensification and broadening of the financial market turmoil is likely to dampen global and euro area demand for a rather protracted period of time. |
| 03/11/2011 | In the Governing Council s assessment, the downside risks to the economic outlook for the euro area are confirmed in an environment of particularly high uncertainty. |
| 07/03/2013 | The GDP outcome for the fourth quarter of 2012 was weak, with Eurostat s second estimate indicating a contraction of 0.6 quarter on quarter. |

## A.3 ECB Statement sentences classification : Topic NONE

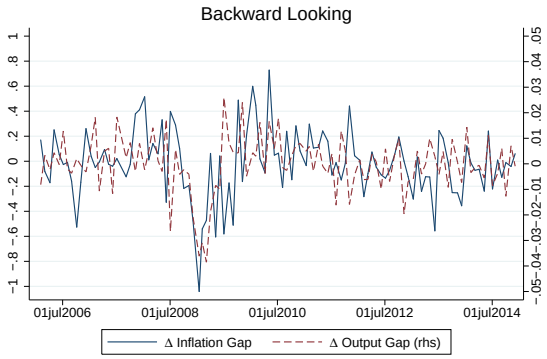
| | |
|---|---|
| **Data sentences** | |
| 09/01/2014 | According to Eurostat s flash estimate, euro area annual HICP inflation was 0.8% in December 2013, compared with 0.9% in November. |
| 06/06/2012 | The June 2012 Eurosystem staff macroeconomic projections for the euro area foresee annual real GDP growth in a range between 0.5% and 0.3% for 2012 and between 0.0% and 2.0% for 2013. |
| 04/04/2012 | The annual growth rate of M3 was 2.8% in February 2012, compared with 2.5% in January. |
| **Repetition** | |
| Each speech | Ladies and gentlemen, the Vice President and I are very pleased to welcome you to our press conference. |
| Each speech | Let me now explain our assessment in greater detail, starting with the economic analysis. |
| +70 times | Over the medium term, inflation expectations remain firmly anchored in line with price stability. |
| 29 times | we will continue to monitor very closely all developments over the period ahead. |
| **Other Topic or explanation** | |
| 08/07/2010 | A lagged response of loans to non financial corporations to developments in economic activity is a normal feature of the business cycle. |

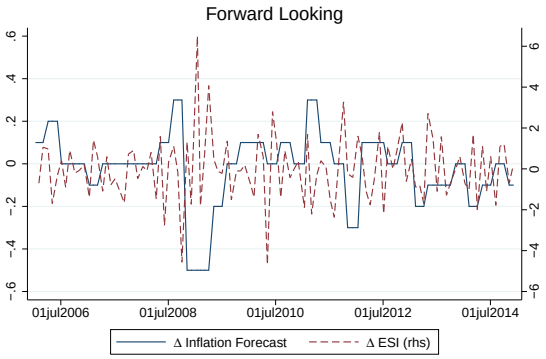
## A.4 ECB Non Standard Policies Announcements

| Date | Reference | Wording from Introductory Statement |
|---|---|---|
| 05/07/2009[1] | CBPP1 | the Governing Council decided today to proceed with its enhanced credit support approach. |
| | 1Y LTRO | we will conduct liquidity-providing longer-term refinancing operations with a maturity of 12 months |
| 08/04/2011 | 6M LTRO | the Governing Council today also decided to conduct a liquidity-providing supplementary longer-term refinancing operation (LTRO) with a maturity of approximately six months |
| 10/06/2011 | LTRO | The Governing Council has decided to conduct two longer term refinancing operations LTROs , one with a maturity of approximately 12 months in October and the other with a maturity of approximately 13 months in December. |
| | CBPP2 | Furthermore, the Governing Council has decided to launch a new covered bond purchase programme CBPP2 . |
| 12/08/2011[1] | 3Y LTRO | First, to conduct two longer-term refinancing operations (LTROs) with a maturity of 36 months and the option of early repayment after one year. |
| | Collaterals | Second, to increase collateral availability by reducing the rating threshold for certain asset-backed securities (ABS). |
| | Reserve Ratio | Third, to reduce the reserve ratio, which is currently 2 , to 1 . |
| 09/06/2012 | OMT | the Governing Council today decided on the modalities for undertaking Outright Monetary Transactions OMTs in secondary markets for sovereign bonds in the euro area. |
| 07/04/2013 | FG | The Governing Council expects the key ECB interest rates to remain at present or lower levels for an extended period of time. |
| 06/05/2014[1] | TLTRO | targeted longer term refinancing operations |
| | SMP | we have decided to suspend the weekly fine tuning operation sterilising the liquidity injected under the Securities Markets Programme. |
| | ABS | preparatory work related to outright purchases of asset backed securities |
| 09/04/2014[1] | ABS | In addition, the Governing Council decided to start purchasing non financial private sector assets. |
| | CBPP3 | the Eurosystem will also purchase a broad portfolio of euro denominated covered bonds issued by MFIs domiciled in the euro area under a new covered bond purchase programme CBPP3 . |
| 01/22/2015 | QE | First, it decided to launch an expanded asset purchase programme, encompassing the existing purchase programmes for asset backed securities and covered bonds. |
| | LTRO | Second, the Governing Council decided to change the pricing of the six remaining targeted longer term refinancing operations TLTROs . |

[1] : The Governing Council also announced an interest rate cut before the Press Conference.

## A.5 Macroeconomic Variables



## A.6 Descriptive Statistics

| | Mean | Std. Deviation | Min. | Max. | ADF t-statistic |
|---|---|---|---|---|---|
| $R_t$ | 1.712 | 1.354 | .05 | 4.25 | -0.127 |
| $\Delta R_t$ | -.016 | .149 | -.75 | .25 | **-6.590** |
| $(\pi_t - \pi^*)$ | -.001 | .010 | -.026 | .021 | -0.955 |
| $\Delta(\pi_t - \pi^*)$ | -.001 | .003 | -.010 | .007 | **-7.928** |
| $(y_t - y^*)$ | .001 | .037 | -.117 | .078 | -1.540 |
| $\Delta(y_t - y^*)$ | -.000 | .011 | -.038 | .026 | **-9.189** |
| $\Delta y_t^e$ | -.0068299 | 1.56758 | -4.679339 | 6.495041 | **-14.150** |
| $\Delta \pi_t^e$ | -.0186916 | .172731 | -.5 | .3 | **-3.794** |
| $\Delta Eurostoxx_t$ | -0.002 | .017 | -.062 | .057 | **-11.344** |
| $\Delta VOL_t$ | -.003 | -.067 | -1.644 | .236 | **-10.057** |

Notes: The ADF test null hypothesis $H_0$ assumes the existence of a unit root. Values in bold reject $H_0$ at a 1% confidence level (the 99% critical value is equal to -3.508) with 0 lag.

Correlation Matrix

| | $\Delta(y_t - y^*)$ | $\Delta(\pi_t - \pi^*)$ | $\Delta y_t^e$ | $\Delta \pi_t^e$ |
|---|---|---|---|---|
| $\Delta(y_t - y^*)$ | 1.0000 | | | |
| $\Delta(\pi_t - \pi^*)$ | 0.2415 | 1.0000 | | |
| $\Delta y_t^e$ | -0.1698 | -0.1729 | 1.0000 | |
| $\Delta \pi_t^e$ | 0.3861 | 0.3718 | -0.2153 | 1.0000 |
| $I_t^{MP}$ | 0.2875 | 0.2564 | -0.1515 | 0.4212 |
| $I_t^{EC}$ | 0.1580 | 0.2501 | -0.1689 | 0.4151 |
| $LM_t$ | 0.2530 | 0.2495 | -0.1973 | 0.5266 |
| $BN_t$ | 0.2308 | 0.3303 | -0.1570 | 0.5059 |

## A.7 Results from Equation 6 using the real time indicators $RT_s^c$

| | Monetary policy decisions $ECB_t$ | |
|---|---|---|
| | $\Delta MRR_t$ | $Decision_t$ |
| $ECB_{t-1}$ | 0.070 | **-0.452*** |
| | (1.174) | (0.231) |
| $\Delta(y_t - y^*)$ | 5.814 | 10.393 |
| | (16.975) | (13.206) |
| $\Delta(\pi_t - \pi^*)$ | 143.538** | 27.389 |
| | (63.102) | (49.622) |
| $I_t^{MP}$ | 0.105 | 0.381 |
| | (0.353) | (0.300) |
| $I_t^{EC}$ | **2.926**** | **2.328***** |
| | (0.726) | (0.559) |
| Observations | 106 | 106 |
| $Pseudo - R^2$ | 0.273 | 0.224 |

Notes: The tables report the results from an ordered probit model estimated with maximum likelihood between January 2006 and December 2014. The dependent variable is, for the upper part, the change of the ECB MRR and, for the lower part, ECB monetary policy decisions. Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

## A.8 Results from Equation 7 using the real time indicators $RT_s^c$

| | Monetary policy decisions $ECB_t$ | | | |
|---|---|---|---|---|
| | $\Delta MRR_{t+m}$ | | $Decision_{t+m}$ | |
| | m=1 | m=2 | m=1 | m=2 |
| $ECB_t$ | 0.134 | 0.779 | **-0.513**** | 0.080 |
| | (1.064) | (0.938) | (0.199) | (0.194) |
| $\Delta(y_t - y^*)$ | 20.003 | 12.249 | 17.162 | **17.842*** |
| | (12.737) | (9.448) | (13.110) | (10.648) |
| $\Delta(\pi_t - \pi^*)$ | 52.870 | -25.277 | -0.513 | -34.431 |
| | (47.592) | (45.222) | (42.365) | (49.586) |
| $I_t^{MP}$ | -0.411 | -0.510 | -0.217 | -0.041 |
| | (0.408) | (0.410) | (0.357) | (0.334) |
| $I_t^{EC}$ | **2.922***** | **2.216***** | **2.979***** | **1.751***** |
| | (0.668) | (0.660) | (0.603) | (0.543) |
| Observations | 106 | 106 | 106 | 106 |
| $Pseudo - R^2$ | 0.263 | 0.170 | 0.244 | 0.154 |

Notes: Coefficients are maximum likelihood estimations of an ordered probit model between January 2006 and December 2014. The dependent variable is the one period ahead ($m = 1$) or two period ahead ($m = 2$) value of, for the upper part, the change of the ECB MRR and, for the lower part, ECB monetary policy decisions . Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

## A.9 Regression results - Market Volatility and Text-Based Market Uncertainty

| | VSTOXX | | |
|---|---|---|---|
| | [1] | [2] | [3] |
| $VOL_{t-1}$ | 0.0866 | 0.1038 | 0.1229 |
| | (0.0906) | (0.0867) | (0.0838) |
| $Surprise_t$ | -0.1966 | -0.1800 | **-0.1832*** |
| | (0.1207) | (0.1094) | (0.1101) |
| $I_t^{MP}$ | **0.0342**** | | |
| | (0.0150) | | |
| $I_t^{EC}$ | **-0.0493*** | | |
| | (0.0292) | | |
| $ULM_t$ | | 1.7542 | |
| | | (1.5764) | |
| $UBB_t$ | | | -0.0086 |
| | | | (0.0118) |
| Observations | 106 | 106 | 106 |
| $Adj. - R^2$ | **0.0519** | 0.0204 | 0.0164 |

Note: The table reports the results from a linear regression of contemporaneous market volatility (Equation 10). Robust standard errors are reported in parenthesis and superscripts ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.