What's new in the news? Media coverage about the ECB and market participants' inflation expectations.

Mikael Eskenazi*

Monday 22nd May, 2017

Abstract

The ECB attracts a lot of media coverage but do these news carry additional information to market participants after they observe economic data and central bank communication? In a framework where an agent uses media coverage about monetary policy to update her prior about the future inflation rate, media coverage affects the mean inflation expectation as well as its variance alongside other traditional information channels. I empirically test these effects using Dynamic Topic Modelling to measure media coverage about the ECB from September 1st 2014 to January 3rd 2016 and prices on 1-year inflation cap options to derive market participants' implied inflation expectations. My results are supportive of a significant positive effect of the changes in the proportion of the ECB media coverage dedicated to the ECB's monetary policy on market participants' mean inflation expectations while the uncertainty in media coverage is found to increase their variance. But the proportion of media coverage dedicated to the ECB's monetary policy is found to increase the variance of inflation expectations.

*I would like to warmly thank Christine Graeff and Benoit Coeure for encouraging this research as well as sharing their expertise and experience during my time at the ECB. I would like to thank the ECB's Global Media Relations division for giving me access to the financial and media data, which were paramount to conducting this work. I would like to especially thank Stefania Secola, Thierry Bracke, Gabriel Glockler, Sakari Suoninen, Martin Zwiebelberg and Michael Steen for their stimulating support and essential comments. I would like to thank Francesco Saraceno for accepting to coordinate my project, for helping me structure my approach and sharpen the research question. Paul Hubert, Stefano Nardelli and an anonymous referee provided invaluable comments. Finally, I would like to thank Olivia for her uncompromising support throughout this research process. Mistakes remain of course solely my responsibility.
Email: mikael.eskenazi@sciencespo.fr

Introduction

It is key to realise that most of the ECB's communications are conveyed by the media. Striving for transparency, accountability and policy effectiveness, the ECB invests a lot of time to communicating with the media. Apart from the most obvious press conference times, where the ECB's President explicitly communicates to a sub-group of specialised journalists, top ECB central bankers participate to hundreds of media activities a year, such as interviews, speeches or background talks. In addition, the ECB has attracted media attention since it released ambitious brand-new policies such as targeted long term refinancing operations, covered bonds purchase programmes or asset purchase programmes, and came to appear at the center stage of today's policy-making. This has translated into a surge in volume of news articles about the ECB, into more commentaries about the ECB's measures and into an overall more complex media coverage about the ECB's monetary policy. But apart from the obvious image management at play, does media coverage about the ECB bring any new information for market participants to forecast inflation after they observe data and the ECB's communication directly?

In an imperfect information setting where an agent has some prior information about the future inflation rate and uses media coverage to update her expectations, the agent's inflation expectation follows a weighted average of her prior information and the media signal she observes. Because of costs to obtaining information, this bayesian learning framework fits consumers' expectations but in the spirit of the literature on rational inattention, market participants could also partly rely on media when collecting information about future inflation because of finite information processing capacities, even at near-zero information costs.
In this paper, I empirically test this framework in order to answer the following research question: does media coverage about the ECB convey additional useful information to market participants for them to forecast inflation alongside the established information channels of economic data and direct central bank communication?
In particular, I test the three following hypotheses : 1/ the volume of media discourse about the ECB's monetary policy affects market participants' mean inflation expectations, 2/ the volume of media discourse about the ECB's monetary policy reduces the variance of market participants' inflation expectations, 3/ the uncertainty in the media discourse about the ECB's monetary policy increases the variance of market participants' inflation expectations.

This paper borrows an automated classification method to the machine learning literature called "topic modelling" in order to measure media coverage in English about the European Central Bank from September 1st 2014 to January 3rd 2016. Specifically, it uses a "dynamic topic model" (DTM) to model semantical evolutions in the media coverage about the ECB. This method extracts *topic proportions* and *word distributions inside topics* from the corpus of press coverage about the ECB, thereby comprehensively addressing how the ECB's monetary policy is being covered in the media. Crucially, DTM explicitly models time dynamics in topics' wording: topics have changing wording over time and are allowed to incorporate new terms as they appear in the data. This allows not only to compute topics' proportions over time consistently, but also the precious time-varying words' distributions inside topics in order to model how changeable media coverage's topics and wording were over the sample.

The first contribution of this paper is to bring a bayesian learning framework, that is often used to model consumers' expectations, to the data of market participants' inflation expectations. I show that although market participants are knowledgeable about monetary policy and are able to gather information directly from the ECB, media coverage of the ECB's monetary policy provides valuable information to their inflation expectation formation. After controlling for the traditional ways market participants receive information from the ECB (data and direct communications), indirect communications in the form of the large media coverage about the ECB are shown to be a significant information channel about future inflation.

The second contribution of this paper to existing literature on central bank communication is a methodological one: I measure media coverage about the ECB in terms of topics' proportions and varying word distributions at a weekly frequency, accounting for more than 40.000 articles in English mentioning the ECB over the sample. To the best of my knowledge, DTM has never been used in applications about central bank communication.

My paper is related to the literature on imperfect information, central bank communication and media studies. In particular, I relate to Carroll (2003) on the role of media in the formation of the level of inflation expectations, but Carroll (2003) focuses solely on consumers' inflation expectations and crucially assumes media only convey rational professional expectations while I focus on market participants' inflation expectations and treat media coverage as an observable variable. On the variance of inflation expectations, my paper is related to Lamla and Lein (2014) and Maag and Lamla (2009) on the role of media coverage about inflation in the accuracy of inflation expectations in a bayesian learning framework. But I apply this framework to market inflation expectations rather than consumers' and, unlike their conclusions, I find that the uncertainty in media coverage also affects the variance of inflation expectations. Furthermore, my result that an increase of volume of media discourse about the ECB's monetary policy increases the variance of market inflation expectations relates to the literature on public information as in Morris and Shin (2012) who predict that under certain circumstances, overreaction to public signals is possible.
My paper also relates to the large literature on the effects of central bank communication on financial markets and particularly to papers that use automated classification approaches in order to measure communication variables, such as Jansen and de Haan (2010) or Nardelli and al. (2017). But these papers focus on measuring media coverage in terms of hawkishness and dovishness while I aim at modelling media coverage in terms of distributions of topics and wording inside topics over time. Hansen and al. (2015) and Hansen and McMahon (2016) also make use of topic variations in communications but focus on FOMC statements and only use a static version of topic modelling called Latent Dirichlet Allocation. Unlike DTM, LDA does not allow topics to have changing wording, which is an issue when considering central bank communication about new policy-related terms such as "quantitive easing" or "negative interest rates".

The key results of this paper are as follows.
Changes in the proportion of media discourse dedicated to the ECB's monetary policy are found to affect the level of market participants' mean 1-year inflation expectations, which confirms hypothesis 1 and is consistent with the theoretical prediction of bayesian learning. In particular, changes in the proportions of media discourse about the ECB's monetary policy and framed in terms of quantitative easing are found to decrease market participants' implied probability of deflation over the one-year horizon. These results hold when controlling for the other traditional information channels of relevant economic data and direct central bank communications.
The proportion of media discourse dedicated to the ECB's monetary policy is found to also affect the variance of market participants' one-year inflation expectations. But unlike what bayesian learning predicts- and with regard to that matter, also Carroll (2003)- an increase of the media signal increases the variance of market inflation expectations.
Lastly, the uncertainty in the media discourse about the ECB is found to increase the variance of market inflation expectations over the one-year horizon, which is consistent with what bayesian learning predicts.

The remaining of this paper is structured as follows. Section 1 describes in more details the literature about the role of media in the formation of expectations I relate to. Section 2 describes the bayesian learning framework in which media coverage affects market inflation expectations and derives three testable hypotheses. Section 3 describes my data and how my dataset on media coverage and inflation expectations is constructed. Section 4 tests the hypotheses and discusses the implications and limitations of the results with regard to the research question. Section 5 displays robustness checks. Section 6 concludes.

1 The role of media in the formation of expectations

My paper relates to various strands of the literature such as imperfect information, central bank communication and media studies.

1.1 Imperfect information and the role of media

First of all, I relate to the literature on imperfect information such as costly information, bayesian learning and rational inattention in the way I set out to analyse the role of qualitative information stemming from media coverage when information is not perfect.

In costly information models, information is not perfect as there are costs to observing it. Mankiw and Reis (2001) model a Philips Curve where only a fraction of agents update their views on the sate of the economy while others keep behaving with respect to outdated information, inducing aggregate sticky prices through optimal pricing decisions based on delayed information rather than exogenous Calvo pricing. Carroll (2003) was the first paper to explicitly model the role of media coverage in the formation of consumers' inflation expectation, in a context where consumer's information set is imperfect. In Carroll (2003), consumers observe professional forecasters' expectations through media news and form their inflation expectations as a weighted average of their past expectation and the rational expectation they observe in media, inducing a lag in how aggregate inflation expectations behave. One key aspect of Carroll's paper is that media are considered to be disseminating rational inflation expectations to consumers without noise, therefore the more media coverage, the more rational consumers' inflation expectations become. But media do not only incorporate professional forecasts nor are they perfectly informative about future inflation, and they could potentially be misleading through conveying the future inflation rate with noise.

Bayesian learning also offers a workable framework for making sense of media coverage in the formation of inflation expectations when information is imperfect. It is the approach that is closest to this paper, although I focus on market participants' inflation expectations while all papers I know of that make use of this framework focus on consumers' inflation expectations. In a context where it is costly for consumers to obtain information about inflation and where they have to rely on news about inflation stemming from the media in order to update their prior belief on future inflation, Maag and Lamla (2009) study the role of media in the disagreement of consumers' inflation expectations. They find that consumers' disagreement is reduced by the proportion of news stories indicating a rising inflation, while they find that the volume of news and their accuracy towards inflation do not have any impact. In a similar framework, Lamla and Lein (2014) study the effect of the German media coverage about inflation on the accuracy of German consumers' inflation expectations and also find that the volume of news stories indicating a rising inflation improves the accuracy of consumers' inflation expectations, but only when media's tone towards rising inflation is neutral, otherwise the effect reverses.
The provision of public information to agents who face a mix of private and public information does not always lead to welfare improvements. Morris and Shin (2012) provide a theoretical framework for the role of public information in such agents' decision-making. Although stemming from game theory, Morris and Shin (2012)'s equilibrium decision is a weighted average of the private signal and the public signal, where the weights depend on the precision of the signals, as in bayesian learning. The public signal, which is made available by media, convey both a signal about the fundamental and acts as a "focal point" for the beliefs about the variable. When agents only observe the public signal, the precision of the public signal decreases disagreement and increases welfare. But when agents observe both the public and the private signal, the relative precision of the signals governs whether the public signal increases or decrease welfare. In particular, when agents' private information is very precise, the increase in the provision of the public signal may give too much weight to noise and lead to overreaction.

These frameworks assume imperfect information arrises from costly information. But Sims (2002, 2006) addresses the issue of imperfect information when it does not occur from costs to

obtaining information but when agents face a constraint on their ability to process information and turn it into actions. Agents' inattention to some information, although accessible at near-zero cost, is called "rational inattention" and has implications for monetary policy as it impacts how people perceive the economy and form expectations. In particular, as shown in Tutino (2011), rational inattention has important implications in areas that are constantly filled with information such as financial markets: given agents' finite information processing capacity, market participants do not keep optimising as continuous information flows, but rather need to allocate some optimal amount of attention to only certain information in order to maximise their action. Hong and al. (2002) show that investors are inattentive to some economic news that should be theoretically taken into account, Huang and Liu (2007) show that investors over-invest or under-invest in some portfolios as they are only able to process information at a limited frequency.
Such a rational inattention to the otherwise cheap full set of information may explain why some readily available information is not used by market participants, or very imperfectly so. Actually, a large empirical literature shows that market participants only use a subset of all freely available information. For example, Veronesi (1999) shows that financial markets overreact to bad news in good times, and discount good news in bad times. Bondt and Thaler (1985) confirms that market participants overreact to surprising news. Sicherman and Al (2015) find that investors pay less attention to the market when markets are low and when market stress is high.
Furthermore, some empirical literature has focused on showing that market participants partly relied on media when making decisions, which may occur even in the context of extremely low information costs with contraints on investors' information processing capability. For instance, Hayo and Neuenkirch (2014) use survey data that show that market participants perceive media coverage of the ECB's communications and decisions as being reliable and tend to rely more often on media reporting than on self-monitoring when it comes to monitor the ECB's interest rate decisions and speeches.

1.2 Central bank communications and the media

My paper is also related to the large literature on the theory and the effects of central bank communication on financial markets as well as how media depict central banks' communications.

Among others, Walsh (2003) shows that when a central bank commits to a specific goal, it can achieve superior results than a central bank that acts with discretion. Woodford (2005), Walsh (2003) and Reis (2013) show that central bank communication has turned into a monetary policy tool in affecting agents' expectations and that with low nominal interest rates, central banks have relied on their communications, such as forward guidance, in order to keep steering longer-term interest rates and maintain the functioning of their transmission mechanism. Blinder and al (2008) reviews most of what there is to know about central bank communication empirics and theory. Authors have found significant impacts of central bank communications on yields and market expectations. In particular, Glick and Leduc (2011) and Bauer and Rudebusch (2013) find that policy announcements "move" key economic variables at least as much as their implementation, Kohn and Sack (2004) study the impacts of chairs' statements, Gurkaynak and al (2004) find that FOMC communications have explanatory power over changes in yields around FOMC events, Ehrmann and Fratzscher (2007b) study the role of ECB speeches, Hubert and al (2017) and Reeves and Sawicki (2007) model the effects of inflation reports on rate and inflation expectations, Hubert (2016) studies the effects of macroeconomic projections, Jung (2016) shows that the publication of the ECB's accounts provides useful information for market participants to predict the next monetary policy decision.

When it comes to the role of media coverage about monetary policy, most research pieces focus on the transmission of the ECB's communication through media favourability and image and less in terms of policy effectiveness. For example, De Haan and al (2004) and Berger and al (2006) study media reporting about the ECB but mostly through media favourability about monetary policy. Berger and al (2006), who are ones of the first to provide a systematic coding of a large dataset of media coverage about the ECB, show that media favourability depends on how policy is communicated. For example, they find that even an unanticipated policy move may lead to favourable media coverage if it is well explained during the press conference. Eskenazi and al (2017) show that the ECB media reputation is different from its favourability in the media as the latter is very short-lived.

Pereira das Neves (2016a,b) are probably one of the first works that explicitly focus on the role of media coverage about monetary policy in the monetary transmission mechanism through expectations. The key conclusion of these papers is that media coverage tends to increase the audience's perception of uncertainty because of the "man-bites-dog" effect. Since tail-probability events make it to media coverage more easily than common events, the media's covering a certain communication leads to an increased perceived volatility. Other works also study the role of media coverage in monetary policy effectiveness: Sturm and Lamla (2013) study the impact of the ECB's monetary policy and communication on interest rate expectations expressed in newspapers and find that the ECB's communication may affect newspapers' expectations. Nardelli and al (2017) study the importance of media reporting in market participants' ability to forecast the ECB's monetary policy decisions and understanding of the ECB's reaction function. Lucca and Trebbi (2011) provide the same analysis applied to the Fed. They show that communication enhances the predictability of monetary policy.

1.3 Communication studies and media effects

I finally relate to communication studies on media framing and agenda-setting and media effects.

The "agenda-setting" role of media were first defined in McCombs and Shaw's seminal article of 1972 about political communication: "in reflecting what candidates say during the campaign, the mass media may well determine the important issues- that is, the media may set the agenda of the campaign". Defined broadly, the media's agenda-setting power means that in producing the media discourse, mass media have the capacity to attach various levels of importance to different issues and have the power to crowd-out topics. Ever since MacCombs and Shaw (1972) identified the media agenda-setting effect by comparing which topics mass media displayed during the election of 1968 and which issue voters declared was important, the effects of mass media were analysed beyond politics.

DellaVigna and Kaplan (2006) analyse the effect of Fox news' coverage of elections between 1996 and 2000 and find that Republicans significantly gained votes in the towns in which Fox News was broadcast. George and Waldfogel (2004) study the impact of the arrival of a new source of information on local turnout to elections and find that turnout decreased where the New-York Times gained reach. When it comes to the effects of media' agenda-setting on financial markets, Engelberg and Parsons (2011) exploit geographical heterogeneity in local media circulation in order to derive causal effects of media coverage on financial markets. In a research about the effect of media's agenda-setting about the Bank of Canada's communications, Hayo and Neuenkirch find that media coverage is more significant when its coverage deviates from the original wording.

2 Model and hypotheses

Following the literature, market participants should have access to large amount of information when forecasting the future inflation rate. Figure 1 summarises the different information channels through which information about future inflation may reach market participants' information set. Market participants are assumed to obtain information about future inflation from quantitative economic data, qualitative central bank communication directly from the central bank as well as qualitative media coverage about monetary policy. Market participants' potentially relying on media coverage to obtain relevant information about the future inflation rate may come from limit on market participants' processing capacity in the spirit of rational inattention models and/or from information cost as in classical costly information models.

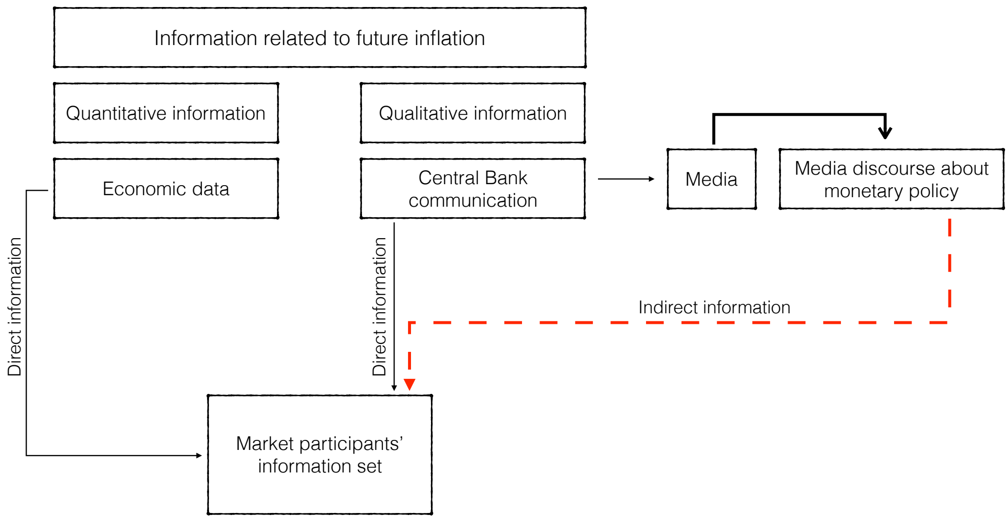

Figure 1: Direct and indirect information in market participants' information set.

I present a simple bayesian learning framework similar to that in Lamla and Lein (2014) and Maag and Lamla (2009), where a representative agent has to forecast future inflation using the Bayes's rule, given some prior knowledge about inflation and an incoming media signal of news. The fundamental derivation of bayesian learning can be found in Jordan (1991) and Evans and Honkapohja (2001).

At the beginning of period t, a representative agent obtain some prior knowledge about the future inflation rate from directly observing a set of relevant economic data of the form

$$P(\pi_t) \sim \mathcal{N}(\mu_{\pi_t}, \sigma_p^2) \quad \text{with pdf} \quad \propto exp\Big\{ - \frac{1}{2}(\pi_t - \mu_{\pi_t})^2 \frac{1}{\sigma_p^2} \Big\}, \tag{1}$$

where σ_p^2 is known with certainty to the agent. During period t, she observes a signal $\xi_{v,t}$ of V news from the media of the form

$$\xi_{v,t} \sim \mathcal{N}(\pi_t, \sigma_s^2)$$

and calculates $f(\xi_{v,t}|\pi_t, \sigma_p^2) = \mathcal{L}(\pi_t, \sigma_s^2; \xi_{1,t}...\xi_{V,t})$, the sample likelihood of the observed media signal under the prior π_t and the known signal's variance σ_s^2. The likelihood function of the media signal is

$$\prod_{v=1}^{V} f(\xi_{v,t}|\pi_t, \sigma_p^2) \propto exp\Big\{ - \frac{1}{2\sigma_p^2} \sum_{v=1}^{V} (\xi_v - \pi_t)^2 \Big\} \tag{2}$$

and is informative of the probability of observing a certain news under different states of future inflation. At the end of the period, the agent updates her prior belief using the Bayes rule and her posterior is of the form

$$P(\pi_{t+1}|\xi_{v,t}) \propto f(\xi_{v,t}|\pi_t) \cdot P(\pi_t) \tag{3}$$

Hamilton (1994) pp. 352, 353 shows that because of the conjugacy assumption from (1) and (2), the product of the type of (3) follows a normal distribution $\mathcal{N}(M^*, V^*)$ where

$$M^* = \frac{\sigma_s^2}{\sigma_p^2 V + \sigma_s^2}\pi_t + \frac{\sigma_P^2 V}{\sigma_p^2 V + \sigma_s^2}\bar{\xi}_t \quad \text{and} \quad V^* = \frac{\sigma_s^2 \sigma_p^2}{\sigma_s^2 + V\sigma_p^2}$$

Therefore, the representative agent's mean inflation expectation is a weighted average of her prior belief π_t and the average signal $\bar{\xi}_t$ she observes in the media, where the weights w are governed by the relative precision of the prior and the observed signal as well as by the volume of media news V:

$$E_t(\pi_{t+1}|\bar{\xi}_t) = w_p \pi_t + w_s \bar{\xi}_t \tag{4}$$

The weight attached to the media signal is $\frac{\sigma_p^2 V}{\sigma_p^2 V + \sigma_s^2} = (1 - w_p) \equiv w_s$ and

$$\frac{\partial w_s}{\partial V} = \frac{\sigma_p^2 \sigma_s^2}{(\sigma_p^2 V + \sigma_s^2)^2} > 0 \tag{5}$$

equation (5) shows that the weight of media coverage in the agent's inflation expectation is an increasing (concave) function of the volume of the media signal. Thus, my first hypothesis is

H1: the media signal affects the level of the agent's mean inflation expectation when its volume increases.

The agent's variance around her mean inflation expectation is given by V^* and depends on the volume of media signal, the variance of the agent's prior and the variance of the signal. Specifically,

$$\frac{\partial V^*}{\partial V} = \frac{-\sigma_p^2 \sigma_s^2 \sigma_p^2}{(\sigma_s^2 + V\sigma_p^2)^2} < 0 \tag{6}$$

which shows that the variance around the mean inflation expectation is a decreasing function of the volume of media signal. Therefore my second hypothesis is

H2: the volume of media signal affects (and reduces) the variance around the agent's mean inflation expectation.

Furthermore,

$$\frac{\partial V^*}{\partial \sigma_s^2} = \frac{\sigma_p^2 V \sigma_p^2}{(\sigma_s^2 + V\sigma_p^2)^2} > 0 \tag{7}$$

which shows that the variance around the mean inflation expectation is an increasing function of the variance of the media signal. Therefore, my third hypothesis is

H3: the uncertainty in the media signal affects (and increases) the variance around the agent's mean inflation expectation.

Next, I describe how my dataset on media coverage and inflation expectations is constructed.

3 Data

3.1 Measuring media coverage about the ECB

DTM is used to extract topic proportions over time and time-varying word distributions inside topics from a large and otherwise unstructured corpus of media articles in English about the ECB from September 1st 2014 to January 3rd 2016, accounting for more than 44.000 articles and 27.000 unique expressions after cleaning. DTM is a class of topic models that builds upon Latent Dirichlet Allocation (LDA), that is presented in Blei and Lafferty (2009).

3.1.1 Probabilistic modelling of media coverage

Today's world is characterised by a surge in the volume of unstructured data. Unstructured data are data that are not readily organised using an explicitly pre-defined model, such as a systematic classification in terms of fixed metadata. It is generally assumed that 85 percent of today's available data are unstructured. This implies that methodologies for retrieving information from those unorganised datasets, classifying data along dimensions of interest to users or simply browsing through immense knowledge bases are not only of the utmost importance to analysts, but are also fundamental for human users to make sense of our common knowledge and for AI to develop efficient representations of large amounts of information.

There has been a growing interest in being able to quantify text over the past years[1]. But most approaches are probabilistic approaches, that treat the observed words as emerging from a data generative process. These methodologies are willing to posit a certain probabilistic generative process in terms of a joint distribution of the observed data and some latent (unobserved) items. Fitting a probabilistic generative model consists in finding the latent items that are the most likely to have generated the observed data, assuming that the generative process is true. This is broadly equivalent to starting from the observed data and "reversing" some assumed generative process in order to recover its unobserved parts.

Topic modelling is a class of hierarchical probabilistic models of text that use multinomial distributions over words to retrieve the hidden topical structure in large corpuses of text. In the context of topic modelling, "recovering the hidden topical structure" means recovering the most likely distributions over words that are assumed to have generated the observed documents' words.
I use DTM's output in order to construct two key variables of media coverage about the ECB: 1/ the proportion of various topics inside the media coverage about the ECB, 2/ the distribution of words over time inside these topics.

As mixed-membership probabilistic models, topic models address the fact that documents are *mixtures of multiple topics* by crucially defining topics as *distributions over a fixed vocabulary of words*: each topic features all the words with different probabilities; differently said, the same words appear in multiple topics with different probabilities. Observed words are treated as observations arising from a probabilistic data generative process that includes an explicitly chained topical structure as latent variables (topic distributions, word assignment to topics, word distribution inside topics). Typically, a topic model specifies how words occur inside topics, which occur inside documents, which occur inside a corpus. The aim of topic modelling is to characterise the conditional distribution of the hidden topical structure's variables on (only) the observed words, which is the posterior distribution. Finally, topic models are unsupervised algorithms: contrary to many approaches in the central bank communication literature, no prior information about the data is required before running the analysis. In particular, the topics are "learnt" by the model without supervision.

Define the dataset formed of D documents $(d_1, ..., d_D)$, that are unique media articles in English about the ECB, where documents consist of "bags of words" $W_d = (w_{d,1}, ..., w_{d,Nd})$[2]. Furthermore, the corpus is divided in weekly time slices $S = (S_1, ..., S_T)$ where S_t is a vector of all documents during one weekly slice t (Monday to Sunday). V is the number of unique words in total across the whole corpus of documents and these V words consist of K topics, where topic $\beta_{t,k} \in \Delta^V$ is a distribution over the fixed vocabulary made of the V unique words in time t. For instance, the V^{th} element $\beta_{t,k}^V$ of topic k in time t is $P(w_v|\beta_{t,k})$. Documents in time t contain all topics K in different proportions $\theta_{t,d} \in \Delta^K$, where the k^{th} element $\theta_{t,d}^k$ of topic proportions inside document d in time t, $\theta_{t,d}$, is $P(\beta_{t,k}|d_d)$.

[1]See Grimmer and Stuart (2012e) or Kalbhenn (2012) for a comprehensive review.

[2]Documents are only considered "bags of words" : the intuition is that topic models only care about the conditional occurrences of words given other words' occurrences, rather than conditional on a specific order. Therefore, the order or words do not matter.

Unlike LDA, in DTM, topics are given the opportunity to change over time, that is, distribution over words may be time-varying[3]. In order to account for distributions over words that change over time, DTM assumes that $\beta_{t,k}$, which is a V-long vector of multinomial natural parameters[4] for topic k in slice t, evolves from the same topic k in slice $t-1$ in the following way,

$$\beta_{t,k}|\beta_{t-1,k} \sim \mathcal{N}(\beta_{t-1,k}, \sigma^2 I)$$
$$P(w|\beta_{t,k}) = exp\Big\{\beta_{t,k} - (1 + \sum_{v=1}^{V-1}(exp(\beta_{t,k,v})))\Big\}$$

which is a state-space model on the natural parameters of the topic multinomial, where σ^2 serves for controlling how changeable topics are (how changeable topic multinomial natural parameters are).[5]

DTM sets the joint distribution of the observed words and the unobserved topic proportions inside documents over time, topic assignment over time and word distributions over time in specifying the statistical process that is assumed to have generated documents in slice t in the following way. If you want to generate a new document, you

(1) Draw topics[6] from $\beta_t|\beta_{t-1} \sim \mathcal{N}(\beta_{t-1}, \sigma^2 I)$
(2) Draw $\alpha_t|\alpha_{t-1} \sim \mathcal{N}(\alpha_{t-1}, \delta^2 I)$ choose a topic distribution over documents

then for each document, you

(a) Draw $\eta \sim \mathcal{N}(\alpha_t, a^2 I)$ choose a topic proportion from the topic distribution over documents

and for each potential word in the document, you

(i) Draw $Z \sim Mult(\pi(\eta))$ choose a topic assignment for each potential words from the topic proportion
(ii) Draw $W_{t,d,n\in V} \sim Mult(\pi(\beta_{t,z}))$ choose a word from the assigned topic's distribution over words

where

$$\pi(\boldsymbol{x}) = \boldsymbol{\gamma} \text{ with each element } \gamma_k = \frac{exp(x_k)}{\sum_j exp(x_j)} \text{ and } \pi(\beta_{t,k})_w = \frac{exp(\beta_{t,k,w})}{\sum_w exp(\beta_{t,k,w})},$$

[3]In LDA, the postulated generative process that fixes dependencies between the observed words and the latent topical structure (the topic proportion per document, the topic assignment and the topics, ie, the word distributions) assumes document interchangeability (documents' joint distribution is invariant to permutation in documents' order) which is too unrealistic. Most corpuses display some clear time structure in how documents occur. For instance, regarding our dataset, it in unlikely that media articles about the Greek crisis should occur before media articles about the launch of the asset purchase programme, because the latter was announced in January 2015 while the Greek crisis happened during the summer of 2015.
Dynamic topic models relax the assumption of document interchangeability and is able to model documents' topical structure in which topics change over time.

[4]The natural parameters of the multinomial are the logs of the probability of each item. Here, the j^{th} element of $\beta_{t,k}$, $\beta_{t,k,j}$, is ln $P(w_j|\beta_{t,k})$.

[5]A logistic normal distribution is then used to map the Gaussian distributions of topic multinomial values $\beta_{t,k}$ to a time-varying probability distributions over words $P(w|\beta_{t,k})$ that is well-behaved.The logistic normal is used as the dynamic alternative to LDA's Dirichlet in order to map sequences of multinomial natural parameters' values (that are unrestricted in all of $\mathbb{R}^V$) to the simplex that sums to 1. Details on this in Aitchison, (1982) and West and Harrison, (1997). Similarly, the hyperparameter α that serves for steering the concentration of topics within documents is allowed to evolve over time, following the motion $\alpha_t|\alpha_{t-1} \sim \mathcal{N}(\alpha_{t-1}, \delta^2 I)$ where δ^2 controls how volatile α is. α_t is also modelled using a logistic normal.

[6]Technically, the topics are pre-existing to the actual documents.

which is a function that maps the unconstrained multinomial natural parameters (as drawn from the Gaussian), to its mean parameters that pertain to the simplex which the word multinomial is drawn from. The simplex is a practical space for representing topics as combinations of words and documents as mixtures of topics because for any point $p = (p_1, ..., p_J)$ in the simplex, $\sum_j p_j = 1$.

Therefore, DTM essentially consists of an ordered sequence of LDA models with some modelling care on how topics and topic proportions should evolve over time. Most notably, it uses a Gaussian state-space formulation and a logistic normal distribution to allow for smooth time-varying topics.
Since the modelling of topic changes occur at the slice level, DTM relaxes the interchangeability assumption amongst slices but documents remain independent within slices. In the context of this paper, this means that DTM models the evolution of topics at the weekly level, but that fitted topics within one week do not account for the ordered occurrence of media articles within this week.
But as shown in Blei and Lafferty (2006), DTM remains more accurate than using LDA over the whole corpus and ordering ex-post the fitted topics at a chosen frequency. As a matter of fact, DTM is not only able to explicitly model the rise and fall of topic proportions at the slice level, but also, DTM's time-varying word distributions address the obviously evolving wording about monetary policy in the media.

3.1.2 Posterior inference

By formulating dependencies that are chained from the observed words to the postulated topics, the convoluted generative process described above allows to think of the collection of documents as emerging from the joint distribution of the unobserved topical structure and the observed words of the form $P(\beta_{t,k}, \theta_{t,d}, z_{n,k,d}, w_{k,t,d})$. But the final goal of DTM is to estimate the individual unobserved $\hat{\beta}_{t,k} = E(\beta_{t,k}|w_{t:T,d:D,1:N})$, $\hat{\theta}_{k,t,d} = E(\theta_{k,t,d}|w_{t:T,d:D,1:N})$ and $\hat{z}_{n,k,t} = E(Z_{n,k,t} = k|w_{t:T,d:D,1:N})$ in order to "measure" media coverage and these can be inferred from the posterior of the general form $P(\beta_{t,k}, \theta_{t,d}, z_{t,d}|w_{t,n,d}) = \frac{P(\beta_{t,k},\theta_{t,d},z_{t,d},w_{t,n,d})}{P(w_{t,n,d})} = \frac{P(\beta_{t,k},\theta_{t,d},z_{t,d},w_{t,n,d})}{\int\int P(\beta_{t,k},\theta_{t,d},z_{t,d},w_{t,n,d})d\beta d\theta}$.
But it is impossible to compute this posterior exactly because of the difficulty to track the last denominator[7]. So one needs to approximate the posterior.

In LDA, β_k is drawn from a Dirichlet that is a conjugate prior to the word multinomial and Gibbs sampling may be used to recover β_k and $\theta_{k,d}$ from the conditional distribution of $z_{n,k}$ on the observed words (integrating β_k and $\theta_{k,d}$ out)[8].

In DTM, because of the non-conjugacy between the Gaussian $P(\beta_{t,k})$ and the multinomial $P(w|\beta_{t,k})$, Gibbs sampling is inapplicable, although Bhadury and Al (2016) derive a special case of Gibbs sampler that is working for DTM. Advanced inference methods must be used in order to approximate the posterior[9].

As proposed by Blei and Lafferty (2006), this paper uses variational inference in order to approximate the posterior[10]. The main intuition behind variational inference is that it works in defining a distribution Q on the posterior's space that is parametrised in order to minimise the distance (the Kullback-Liebler divergence) to the true posterior. The resulting distributions in Q may be used as an approximation of the true posterior and its optimised variational parameters can be used for computing their true posterior counterparts.
Blei and Lafferty (2006, 2009) show that DTM's approximate variational posterior is given by

[7]Due to the exponentially large dependencies between latent variables in the topical structure when conditioned on the observed documents. See Blei and Lafferty (2009, 2012).

[8]Griffiths and Steyvers (2004)

[9]Also a Metropolis-Hasting algorithm in Bhadury and Al (2016), variational inference in Blei and Lafferty (2006).

[10]I adapt and make use of David Blei and Sean Gerrish's `C++` implementation of variational inference for DTM, which is the same that underlies Blei and Lafferty (2006).

$$Q(\beta_{k,t}, \theta_{t,d}, z_{t,d,n}) = \prod_{k=1}^{K} Q(\beta_{k,1}, ..., \beta_{k,T} | \hat{\beta}_{k,1}, ..., \hat{\beta}_{k,T}) \times \prod_{t=1}^{T} \left(\prod_{d=1}^{d_t} Q(\theta_{t,d}|\gamma_{t,d}) \prod_{n=1}^{n_{t,d}} Q(z_{t,d,n}|\phi_{t,d,n}) \right)$$

which reflects the chained dependencies postulated in the generative process described above and where the sequence $(\hat{\beta}_{k,1}, ..., \hat{\beta}_{k,T})$, the distributions $\gamma_{t,d}$ and $\phi_{t,d,n}$ are the free variational parameters for the sequenced topics $(\beta_{k,1}, ..., \beta_{k,T})$, the topic proportions $\theta_{t,d}$ and the topic assignment $z_{t,d,n}$ respectively. Crucially, the variational distributions of the latent variables are now independent while conditioned on the data in the true posterior, they are highly dependent- which was the issue in computing the posterior- while retaining time dependency between the $\beta_{k,t}$.

The problem is now to find the variational parameters that minimise the distance between $Q(\beta_{k,t}, \theta_{t,d}, z_{t,d,n})$ and $P(\beta_{t,k}, \theta_{t,d}, z_{t,d}|w_{t,n,d})$, that is

$$\arg \min_{\gamma_{t,d}, \phi_{t,d,n}, (\hat{\beta}_{k,1}, ..., \hat{\beta}_{k,T})} \mathrm{KL} \Big(Q(\beta_{k,t}, \theta_{t,d}, z_{t,d,n}) || P(\beta_{t,k}, \theta_{t,d}, z_{t,d}|w_{t,n,d}) \Big).$$

Variational inference may be described as the numerical method for maximising the objective function of this problem with respect to the variational parameters. The algorithm works in updating each variational parameters in the objective with their expectation of their true posterior counterparts under the variational distributions, until the objective (that is in a form of a log likelihood really) converges[11]. The maximised variational distributions can now be used to compute the expected topic proportions and word distributions.

3.1.3 Implementation of DTM

Variational inference finds a local maximum of the objective function that minimises the distance between the variational posterior and the "true" posterior, thereby "learning" the topics in an iterative process. It is worth emphasising that this learning process occurs without any supervision or prior information about the documents, and uses only the observed word counts. Specifically, the objective function is in a form close to a likelihood of the data, in terms of the sample counterparts of the latent topical structure, for example, in terms of $E(\ln P(\beta_{k,t,w}))$ in place of the topics' distribution over words[12].
Therefore, one key element in implementing DTM is to represent the data in terms of particular objects, so that the algorithm is amenable to using the observed words. My dataset is a list of daily full media articles in English about the ECB, from September 1st 2014 to January 3rd 2016, accounting for 70 weekly slices of a total of 43.961 unique documents and 569.757 unique words.

My approach about selecting media outlets in my sample is rather agnostic: I simply collect *every* articles in English where the word "ECB" is mentioned. Specifically, I don't choose any particular outlets prior to the analysis nor do I weight outlets with regard to their reach. I am interested in the overall media semantics about the ECB, so I would rather observe everything than imposing some exogenous reading grid (also, reach is hard to measure). In a way, I simply let the observed volume of each outlet's articles weight their semantics in my sample.
There are 326 unique media outlets in my sample, but only 17 media outlets reach 1% or more of total coverage individually and account for 70% of total coverage collectively. Since I focus on media coverage in English about the ECB, it is clear that most outlets are either UK or US located, but most of them have international circulation or European versions. Furthermore, all

[11] Details on the algorithm in Blei and Lafferty (2006), on variational inference methods in Jordan and al, (1999) in Blei and al, (2003) and in Wainwright and Jordan, (2005).

[12] This is initialised by choosing some random documents. Smoothing their aggregated word counts over the vocabulary creates a first distribution over words, from which $E(\ln P(\beta_{k,t,w}))$ is approximated.

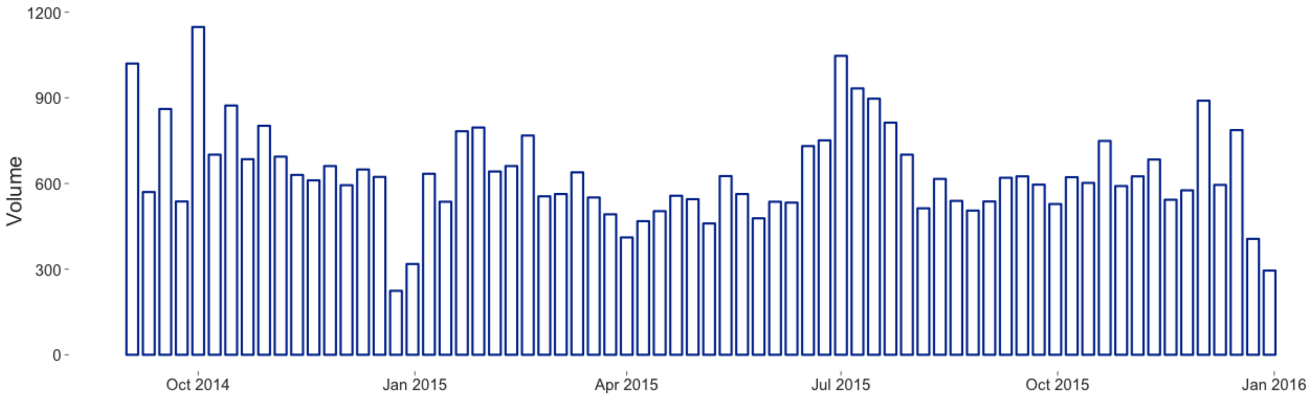

Figure 2: Number of weekly media articles written in English about the ECB over the sample.

of them address business customers and are relevant for market readers. Finally, all of them are highly respected outlets and their content - at least in terms of topics - generally flows to other smaller outlets. The reason why the Irish press gathers a significant proportion in my sample is because it covers banking supervision topics with high volumes, but since I focus on articles about the ECB that covers monetary policy topics, such articles are mostly drawn away from the final analysis.

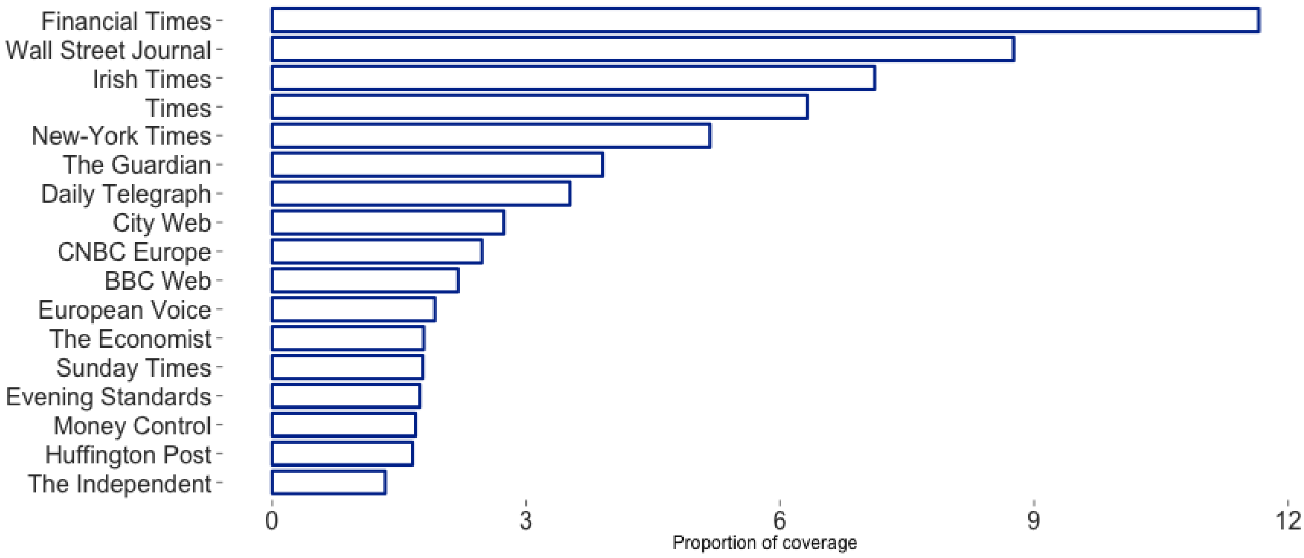

Figure 3: Media outlets representing more than 1% of total coverage in the sample.

This is equivalent to an average of 635 media articles written (in English only!) about the ECB on a weekly basis over the period. Moreover, I calculate that on average, media coverage about the ECB increases by 21 percent on press conference weeks.
In order to explore this corpus and estimate the topic proportions over time (what the media talk about when writing about the ECB) and the time-varying word distributions (how the media talk about certain ECB-related topics), I estimate a 10-topic DTM[13].

In natural language processing, the usual way of representing a corpus of text is to transform it into a Document-Term-Matrix (dtm) A, that summarises the occurrences of the unique words

[13]The number of topics k to be estimated is indeed to be decided by the analyst as this represents the dimension of the simplex, from which the word multinomials are drawn from. Fitting a 20-topic DTM or a 50-topic DTM does not modify the estimated distributions crucially, but is much more demanding in terms of computational power and time.

in the corpus' documents in the form of

$$\underset{(D\times V)}{A} = \begin{bmatrix} freq(w_1|d_1) & \dots & freq(w_V|d_1) \\ \vdots & \ddots & \\ freq(w_1|d_D) & \dots & freq(w_V|d_D) \end{bmatrix}$$

where, in the case of DTM, documents $(d_1, ..., d_D)$ are also ordered in time. But it is clear that all words should not enter A. First of all, because the size of A would be too large. Mostly because all words do not convey purposeful semantical meaning, such as stop words (but, and, the, ...etc). Lastly, because I aim at obtaining topics that are easy to interpret and to label- and that clearly convey semantics inside documents- one needs to select the vocabulary DTM will be estimated with. The selection goes as follows[14].

(1) **Tokenisation**: the vector of documents is transformed into a corpus and documents are cleaned from strange characters, punctuation, numbers, email addresses, webpages and stop words (A = 44.418 documents $\times$ 533.693 unique terms).

(2) **Forming expressions**: the remaining terms of a minimum length of 3 characters are transformed into bi-grams and tri-grams that are expressions formed by all combinations of two and three words (A = 44.418 documents $\times$ 17.120.952 unique terms).

(3) **Filtering terms**: only terms occurring at least in 1/1000 documents are kept (A = 44.418 documents $\times$ 27.104 unique terms).

(4) **Deleting empty documents**: filtering items produces empty documents in A and these are deleted without modifying the time structure in the order of the documents (A = 43.961 documents $\times$ 27.104 unique terms).

Some more objects are created from this final A, and stored,

$$\underset{(27.104\times 1)}{W} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{27104} \end{bmatrix}, \underset{(71\times 1)}{S} = \begin{bmatrix} 70 \\ S_1 \\ \vdots \\ S_{70} \end{bmatrix} \text{ and "count"} = \begin{bmatrix} index_i : count_i|d_1 & \dots & index_{N,1} : count_{N,1}|d_1 \\ index_j : count_j|d_2 & \dots & index_{N,2} : count_{N,2}|d_2 \\ \vdots & \ddots & \\ index_k : count_k|d_D & \dots & index_{N,D} : count_{N,D}|d_D \end{bmatrix}$$

where W represents the total vocabulary and in which each term is given an index equal to its row; where S represents the time slices in which S_t is the number of documents in slice t and whose first entry is the total number of slices; and where "count" is a file listing which terms appear in which documents and their respective counts[15]. The resulting processed dataset is summarised in the table below and the constructed files W, S and "count" are amenable to be used by DTM.

Table 1: Final dataset after cleaning.

Type	DTM variable name	Number of observations
Unique documents	D	43.961
Unique words (before cleaning)	W	569.757
Unique tokens (after cleaning)	V	27.104
Weekly time slices	S	70
Topics	K	10

[14] `R` is used for processing text.

[15] Different indexes are used in order to emphasise that documents do not necessarily contain the same words in the same order.

3.1.4 DTM results

Posterior inference takes approximatively 6 hours on a 3,1 GHz Macbook Pro laptop; and after convergence, I am provided with the individual time-varying word distributions inside topics $\beta_{k,t}$ and the topic proportions over time $\theta_{d,t}$. I only focus on presenting the topics of interest to this paper. Some are illustrated below by their five most probable words.

Table 2: Examples of topics $\beta_{k,t}$

$\beta_{0,1}$	$\beta_{5,1}$	$\beta_{8,1}$
prime minister	european central bank	finance minister
yes vote	mario draghi	prime minister
european commission	quantitative easing	angela merkel
european union	interest rates	international monetary fund
member states	asset backed securities	mr hollande

Table 2 illustrates three different topics in time slice 1, i.e, from September 1st to 7th. It is intuitive to see that because they are distribution over words, when one orders $\beta_{K=k,T=t}$ by most probable words, "topics" are clusters of expressions that make sense together. For instance, it is clear that $\beta_{5,1}$ is about the ECB's monetary policy.
In traditional LDA, it is generally the case that one labels β_k by the expression that is the most likely to occur within it. Since it has time-varying word distributions, DTM makes it harder (or simply, time consuming) to interpret and label topics. Table 3 below illustrates how $\beta_{5,t}$ evolves over time.

Table 3: Some instances of $\beta_{5,t}$

$\beta_{5,5}$	$\beta_{5,20}$	$\beta_{5,65}$
european central bank	european central bank	european central bank
mario draghi	quantitative easing	mario draghi
asset backed securities	government bonds	quantitative easing
balance sheet	monetary policy	deposit rate
covered bond	european court justice	cut deposit

Time slice 5 corresponds to September 29th to October 5th, 2014. Time slice 20 corresponds to January 12th to January 18th, 2015. Time slice 65 corresponds to November 30th to December 6th, 2015. So one also has to appreciate how one topic evolves over time in order to tell what it is about. In this case, it is pretty clear that $\beta_{5,t}$ is about the ECB's monetary policy.
Therefore, using $\theta_{d,t}$ for topic 5, one may quantify how much the topic about the ECB's monetary policy was written about in the media coverage about the ECB over time. Also, from $\beta_{5,t}$, one may quantify how the media's wording about the ECB's monetary policy evolved over time.

A similar inspection is carried for the 10 individual topics $\beta_{1:K,1:T}$ and 5 topics of interest are labelled in table 4 below[16].

DTM is indeed an extremely appealing tool for quantifying unstructured text such as the media coverage about the ECB. Despite the baroque process it assumes words are generated from, DTM seems to be able to gather words into topics that intuitively make sense. Not only do topics form static lists of words that provide the user with meaningful themes, but

[16] Other topics are omitted because they are less obvious to label. Also, some of them are "metadata" topics. Indeed, the original text documents always incorporate an ordered preamble with the name of the author, the outlet and the date. While the date is filtered away by the cleaning procedure described above, some metadata-related terms do remain in the document-term-matrix. Because of their recurrent co-occurring structure, they occur to form "topics"- which I discard. Lastly, some are "data" topics that only contain economic data the media often quote and comment on- I also omit them. Bigger parts of the topics' word distributions are shown in the appendix.

Table 4: Main topics' labels

$\beta_{k,t}$	Label
$\beta_{0,t}$	UK politics
$\beta_{1,t}$	The Fed's monetary policy
$\beta_{2,t}$	The international economy
$\beta_{5,t}$	The ECB's monetary policy
$\beta_{8,t}$	EU politics

also DTM constructs dynamic topics whose time-varying most likely semantics reveal themes that are coherent over time. Below is an intuitive example of how DTM can be applied to one particular document. Figure 2 shows how it is possible to reconstruct a document using DTM's output parameters, $\theta_{d,k}$, $\beta_{K,T=1}$, $z_{n,k,t}$. It also illustrates the intuition behind why it is possible to deconstruct a document using DTM's parameters.

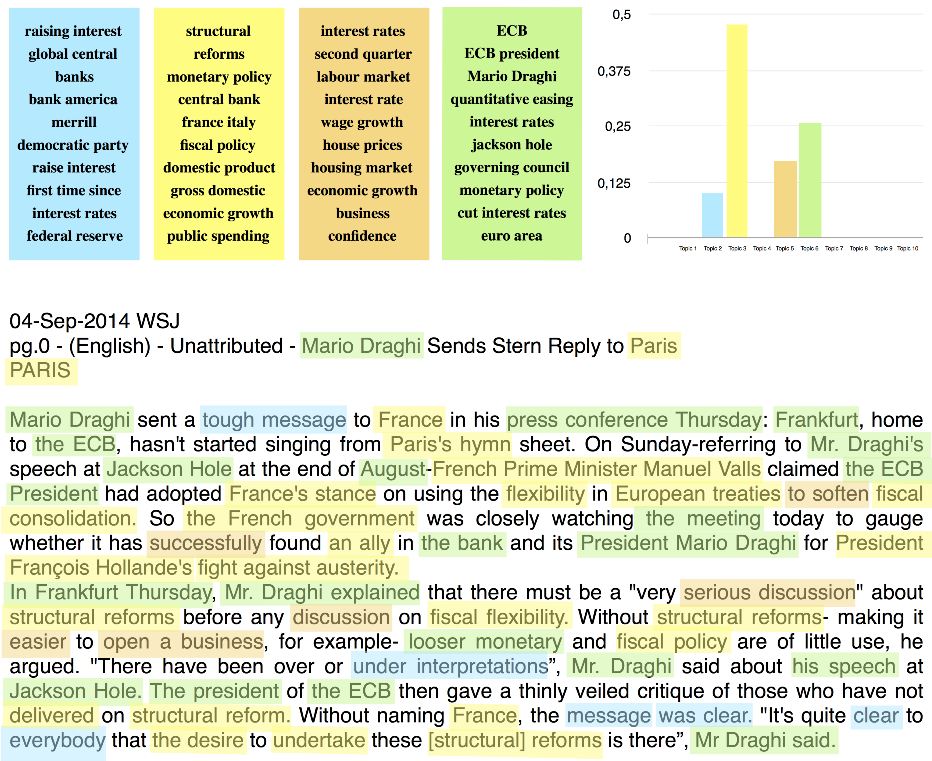

Figure 4: Example of the analysis of one document using DTM's outputs. Topic proportions (top-right corner) are computed from $\theta_{d,k}$. Topics' most likely terms (rectangles in the top-left corner) are taken from the corresponding $\beta_{K,T=1}$. September 4th belongs to the first time slice. Coloured words in the articles are topic assignments $z_{n,k,t}$.

3.1.4.1 Topic proportions over time

By aggregating $\theta_{d,k,t}$ over all the documents in each time slice t, one can compute the topic proportions over time. The proportion of topic k in slice t is defined as $\frac{1}{S_t}\sum_d \theta_{d,k,t}$ and provides an overview of what the media have talked about when talking about the ECB.

Topics' volumes are coherent over time, and extremely consistent with underlying events-with key topics gaining steam at key important topic-related dates. For instance, while media articles about the ECB were mostly about the ECB's monetary policy in the beginning of the sample, the ECB's media coverage is marked by a tremendous rise of the EU-politics-related topic during the summer of 2015. Some instances of the latter topic are displayed in table 5 and it is clear that "EU politics" is mostly about the Greek crisis. It is therefore no surprise that it should peak in March 2015 and during the summer of 2015, when such tensions occurred.

Table 5: Some instances of "EU politics" in 2015

12-18 January	23-29 March	10-16 August
mr tsipras	alexis tsipras	european central bank
prime minister	angela merkel	bailout deal
syriza party	greek banks	debt relief
general election	greek government	finance ministers
central bank	finance minister	new bailout

The topic proportions in media coverage about the ECB display some clear dynamics over time. Topic proportions display some volatility but it is easy to discern trends in rise and fall of certain topics.

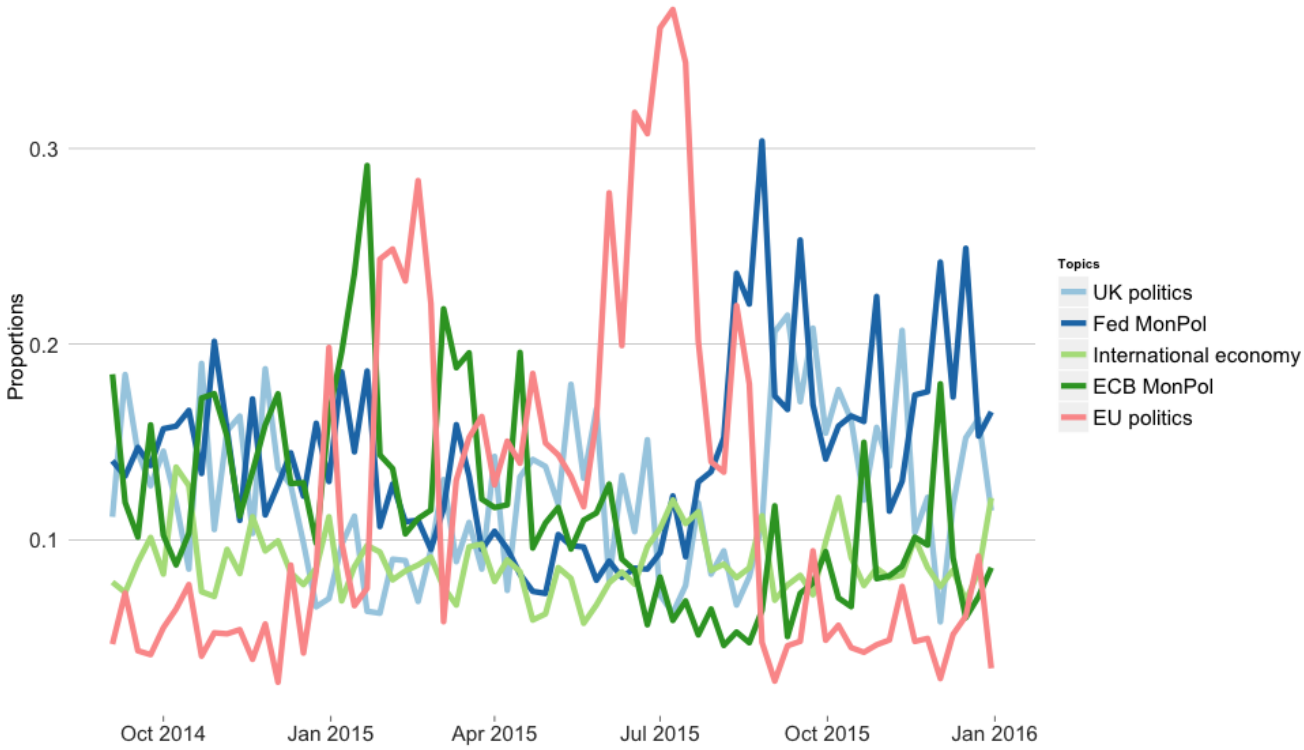

Figure 5: Main topics' proportions over time.

The above figure plots topic proportions over time. It shows how well DTM was able to model topic changes over time, and illustrates how DTM can be used for information retrieval. First of all, this figure provides a quick and dynamic overview of the topics the media have concentrated on when writing about the ECB over the period. It is striking to see how "little" the ECB is being talked about with regards to its monetary policy, relatively to the other topics. Since I use only articles in English that are mostly written in the international press, it is not surprising that the ECB should be often written about with regards to British topics ("UK politics") or other central banks' monetary policy ("Fed's monetary policy").
This confirms the fact built by experience and cross-analyses that the ECB is often portrayed by the media using other topics than core monetary policy, such as politics.

3.1.4.2 Time-varying word distributions

This section plots the $\beta_{k,t}$ for a selection of most likely words and topics. Time-varying word distributions provide a dynamical portrayal of how media have been talking about certain ECB-related topics.

One aspect of the word distributions over time is that they are very large objects. Namely, they consist of tables of as many rows as the number of unique expressions in the vocabulary and as many columns as the number of time slices. In our case, each topic $\beta_{k,t}$ is a table of 27.104 rows and 70 columns. From a human perspective, $\beta_{k,t}$ may be used for information retrieval, in order to recover and quantify each topic's evolving wording. Therefore, representing the time-varying distributions over words is mostly a matter of exploration and it is natural that I present only a small subset of each topic's word distributions. Namely, I extract the 5 to 10 most likely expressions to appear inside a topic from exploring each topic at the slice level. Therefore the interpretation of the extracted time-varying word distributions inside topics goes a follows: they represent how likely over time the media are to pick one particular expression when writing about one particular ECB-related topic.

While figure 1 showed that the media often portrayed the ECB using political topics, "zooming" into the "EU politics" topics shows that the media have also associated such a topic to the ECB. It is indeed clear that the expressions "European Central Bank" or "Mario Draghi" are featured inside the "EU politics" topic with high probability, particularly during the summer of 2015. While the overall media wording of the "EU politics" topic leans towards "Chancellor Merkel" on average, it is during March and the summer of 2015 that the media wording about "EU politics" switched to ECB-related terms.
Looking at the Fed's most likely wording, it is clear that the media have described the Fed's policy in terms of rate hikes and rate cuts with obvious seasonality. Also, there are some indications that they have started to write about the Fed's monetary policy in terms of a rate hike after the summer of 2015.
When picked up in relation to the ECB, the "UK politics" topic was mostly talked about in terms of the Brexit referendum. There are some indications that the media have framed the topic in terms of the "leave EU" expression with some upward trend over the sample[17].

Lastly, it is obvious that the media frame the topic of the ECB's monetary policy in terms of the ECB's various policies over time. For instance, in the beginning of the sample, the media wording about the ECB's policy is focusing on "cut interest rates" and "asset-backed securities", while the end part of the sample is driven by wording on "cut deposit rate". On average, media frame the ECB's monetary policy in terms of its "Quantitative Easing", which is consistent with the period under study[18].

3.2 Computing market inflation expectations from prices on 1-year cap options

Since financial products generally bring revenues in the future, virtually all financial assets' valuation encapsulate some degree of inflation compensation. Some assets are explicitly backed on

[17]The topics on "Fed MonPol" and "UK politics" are briefly discussed as they are not core to this paper's interest in the media coverage of the ECB's monetary policy.

[18]First of all, this shows how finely DTM is able to track wording over time in the context of evolving topics and therefore, why it can be used in order to efficiently quantify unstructured data such as how media present topics on monetary policy and how changeable such a presentation may be. In this regard, DTM is superior to all supervised classifications that work within pre-defined coding categories because DTM is able to model new documents with new topics (that is, topics that were not already decided before coding) and new wording about topics, without losing the advantage of time comparability. DTM's time-varying word distributions $\beta_{k,t}$ can also be used to initialise supervised algorithms, such as these of the hawkish/dovish type, with coherent representations of wordings that change over time. One key challenge of supervised text classification is indeed that words to portray things like monetary policy truly are time-varying, and that fixed reference documents often do not contain sufficient information to classify popping new concepts, such as quantitative easing or negative interest rates.

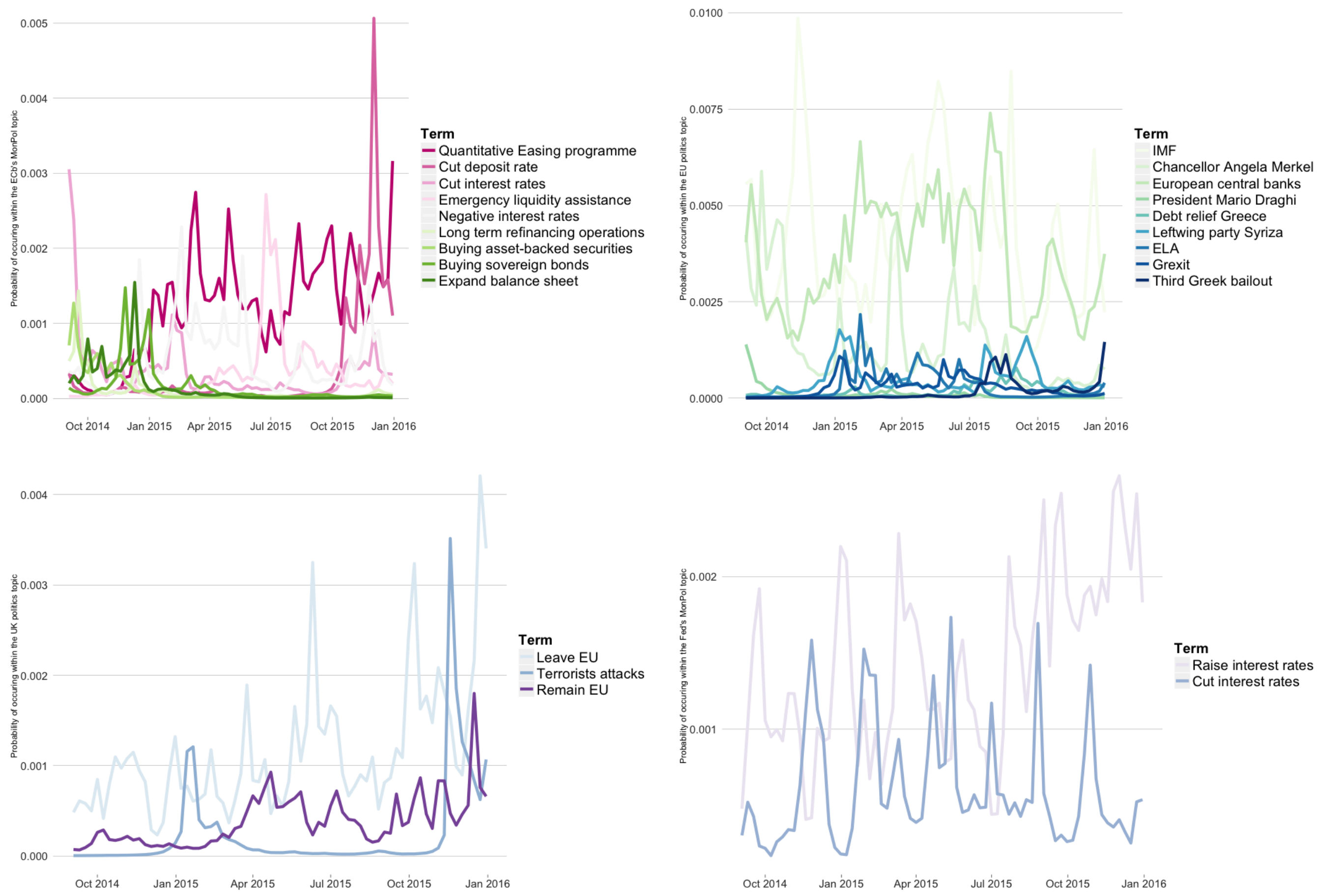

Figure 6: Time-varying word distributions

future inflation rates and may be used to extract market participants' inflation expectations.
Most central bank communication literature has focused on professional forecasters surveys or inflation-linked swaps in order to derive inflation expectations[19]. While the former provides low frequency explicit inflation expectations, the latter allows to derive high frequency implicit expectations- but both surveys and swaps provide point estimates, that is, generally, mean inflation expectations.
However, it is more useful to derive inflation expectations' full distributions rather than sole point estimates, as inflation expectations' higher moments may be calculated from such distributions. For instance, the variance should inform us on the level of disagreement or of uncertainty among market participants about the future inflation rate. Such distributions may be constructed from the observed prices of inflation options, but up to now, relatively little attention has been given to using these data in order to infer inflation expectations- although interest has been growing rapidly[20].
This paper uses prices of European 1-year options in order to derive risk-neutral implied probability density functions for inflation expectations over the one-year horizon, and calculates variance in expectations as well as probabilities assigned by market participants to future deflation and high inflation rates. While there has been recent research on providing researchers with procedures for deriving implied densities from option prices[21], available techniques usually break apart with negative interest rates and need to be adapted to markets' new conventions.

There is a simple idea behind using inflation option prices for constructing distributions of inflation expectations. When giving a price to an inflation option- that gives the right but not

[19]See for instance Hubert (2016).
[20]For example, Smith (2012).
[21]See Bliss and Panigirtzoglou (2000), Andersen and Wagener (2002), Malz (2014), Blake and Rule (2015).

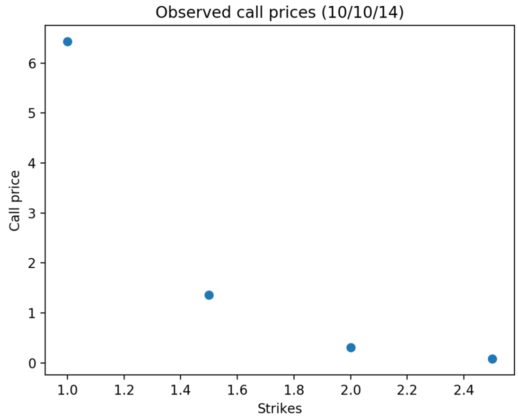

Figure 7: Observed 1-year call prices on 10/10/14.

the obligation to receive a pre-defined inflation rate if the observed inflation rate in the future is above a certain strike rate- market participants should incorporate some subjective probability that inflation will indeed be above the strike rate[22]. Therefore, using option prices at different strike prices but for the same expiration date, one should be able to recover market participants' implied probabilities given to the underlying inflation rate at each strike price and at a certain maturity. The relationship between the implied probability of the underlying inflation rate and the observed cap price is given by Cox and Ross (1976) as

$$C(t,T,K) = e^{-r(T-t)} \int_K^{\infty} w(S_t)(S_t - K) dS_t$$

which relates the price of a call option in time t, at a strike price K expiring at T $C(t,T,K)$ to the density probability function of the underlying $w(S_t)$. But Breeden and Litzenberger (1978) observe that

$$\frac{\partial^2 C(t,T,K)}{\partial K^2} = e^{-r(T-t)} w(K) \tag{8}$$

that is, twice differentiating the call price function $C(t,T,K)$ with respect to the strike price K yields the implied probability density function of the strike price K (up to a discount rate)- which is what we want to derive[23]. However, the call-price function, that is the function that relates the price of an option to various strike prices, is intrinsically discrete because option prices are only observed at precise strike prices. Moreover, only a few strikes are actually observed because cap options at strike prices that are too far away from the forward rate are not traded. Therefore, the main challenge in computing implied inflation expectations' density functions lies in constructing continuous call-price functions in order to apply (3).

Some points are note-worthy about the call prices displayed, and justify my computational approach. First of all, each point represents the value of the 1-year option at a particular strike, and one does not observe any price at strikes below 1%, but one could certainly be interested in market participants' deflation expectations in late 2014: so one needs to extrapolate the call-price function. Secondly, it is clear that call prices are observed in discrete steps since there is no option that is traded at strikes such as 1.80%, but we need a continuous call-price function in order to apply (1): so one needs to interpolate the call-price function. Finally, the call price

[22]For simplicity, I only describe the functioning of a cap inflation option, i.e, an inflation that pays off if the observed inflation rate is above the strike price. The functioning of a floor option would be analogous.

[23]It is important to note that these implied probability density functions will be risk-neutral. Therefore, they should capture market participants' inflation expectations and/or term premia, such as risk premia or liquidity premia.

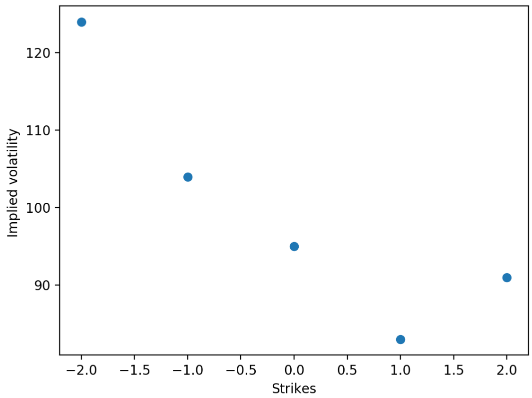

Figure 8: Observed 1-year implied volatilities on 12/12/14. Implied volatilities tend to form a "smile" with a minimum at the at-the-money strike, that is, the strike corresponding to the forward inflation rate.

function exhibits its usual shape: strikes that are above the underlying's price (the forward inflation rate here), out-of-the money strikes, are cheaper than in-the-money strikes that are below the forward rate, which have more value since they are more likely to be exercised.

In order to extrapolate the call-price function to far-out-of-the-money (high inflation strikes) and far in-the money (negative strikes) unobserved strike prices and to interpolate between strike prices, one technique relies on a stochastic volatility model called the SABR[24]. It is easier to fit the call-price function in the implied volatility/strike space, called the volatility smile. The volatility smile is a pricing and trading framework widely used in applied finance, and relates options' strikes to their implied volatilities, instead of their price. Implied volatilities are a measure of call price's sensitivity to the price of the underlying (here, the forward inflation rate). The volatility smile of the 1-year 12/12/14 cap option is displayed below.

The SABR is a technique for modelling this volatility smile, and therefore can be used to extrapolate/interpolate the call-price function by converting continuous volatilities back into continuous prices using any pricing model, for example the Black-Scholes model.

This well-rounded technique is problematic in the context of negative rates. First of all, with negative rates, the whole term structure of inflation expectation shifts downwards, particularly in the short-run. Therefore, negative rates imply near-zero forward rates. Secondly, since implied densities are centred at the forward rate, some parts of the inflation expectations' distributions will lie around negative strikes.
The SABR is derived from the Black's model, which models the forward rate as a logarithmic variable. Therefore, the SABR is incapable of modelling implied volatilities around negative strikes. Furthermore, converting the fitted implied volatilities back into prices also requires a pricing device that accepts negative strikes and forward rates that are possibly negative at some points over the period.

This paper employs a shifted-SABR in order to fit the volatility smile with negative strikes and a shifted-Black model to convert these back into continuous prices[25]. It is crucial to understand that both the SABR and the Black models are needed as fitting device, in order to approximate the volatility smile in a continuous manner and to convert the estimated volatilities back into prices. At no time it is assumed that the Black model is the right pricing model for inflation options, i.e, there is no assumption formed on the behaviour of the forward rate. Therefore, one may just shift the forward rate and the strike prices by some constant in the SABR and in the Black model, in order to model negative entries.

[24] See the original paper on the SABR Bartlett (2006).
[25] The formulas for the shifted SABR and the shifted Black model are in appendix.

Using Bloomberg data on weekly European 1-year inflation cap options from September 1st, 2014 to January 3rd 2016[26], I therefore construct implied inflation expectations' density functions in the following way:

(1) Observed discrete prices are converted into discrete implied volatilities. (2) I calibrate the *shifted* SABR on a range of positive strikes near the at-the-money strike (typically, strikes for calibration include [0%, 0.5%, 1%, 1.5%]). (3) I extend the fitted volatility smile to the negative strikes and to the positive strikes far out of the money using the calibrated parameters (the final support is [-3% ; 4%]). (4) Continuous implied volatilities are converted back into continuous prices using the shifted Black model. (5) The estimated continuous call-price function is twice-differentiated (using finite differences). Figure 4 below summarises the procedure.

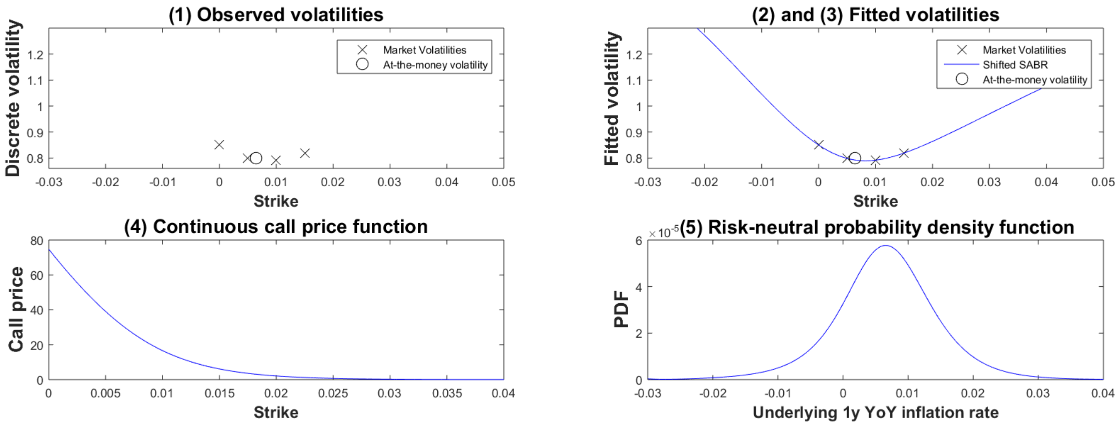

Figure 9: Procedure for computing inflation expectations' implied densities.

Looping over this procedure for each vector of weekly observations of inflation option, and aggregating the resulting densities into a surface yields the following time-varying inflation expectations' risk-neutral implied density function:

[26]I use weekly inflation expectation data because Hayo and Neuenkirch (2014) surveys that market participants perceive the ECB's communication to persist and affect data for most likely one week.

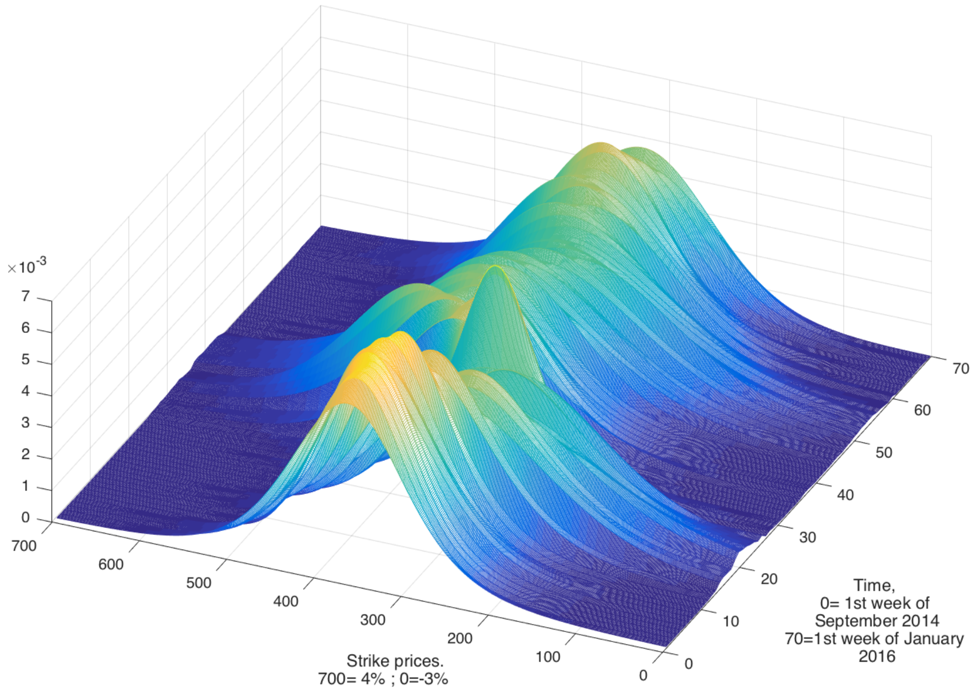

Figure 10: Distributions of 1y inflation expectations implied from option prices from September 1st, 2014 to January 3rd, 2016.

Finally, I use these distributions to calculate the implied expectations' variance and the implied probability of deflation/high inflation over the period.

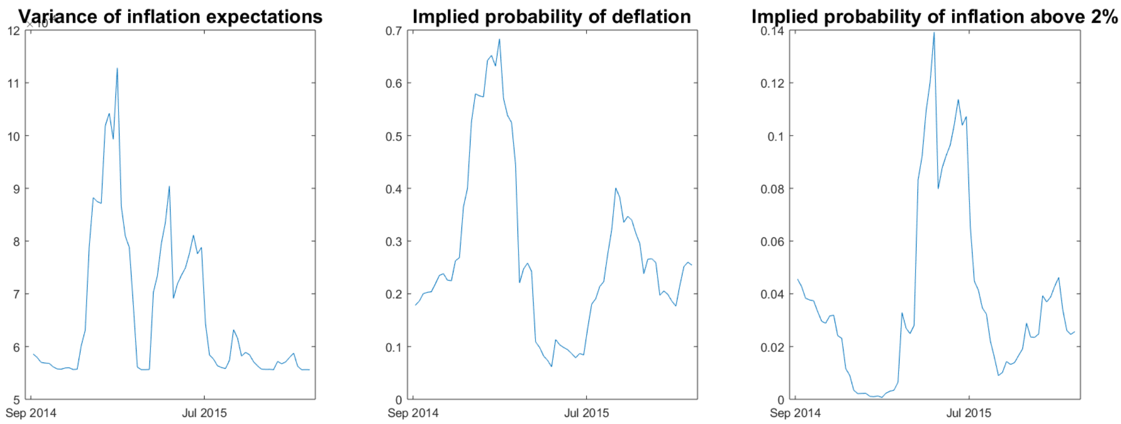

Figure 11: Variance, probability of deflation and probability of high inflation implied by the risk-neutral densities.

3.3 Other data

In order to avoid omitted variable bias in estimating the effect of media coverage on market inflation expectations, I need to control for the two other information channels market participants may receive information about the future inflation rate from. The first channel goes through observing a set of relevant economic data. The second channel goes through observing central bank communication directly. With regard to the model presented in section 2, these vectors

Table 6: Descriptive statistics for inflation expectation data

	N	min	max	mean	sd
Mean inflation expectation	70	-.0046941	.0117	0.00427	0.00400
Variance inflation expectation	70	.0072449	.0128672	0.00834	0.00142
Probability of deflation	70	.1091786	.6759329	0.316	0.149

can be thought of a set of observables market participants use to construct some prior belief about future inflation.

When it comes to inflation mean expectation, I include in the set of relevant economic data market participants ought to observe the level of CPI, the level of conventional refinancing operations the ECB conducts on a weekly basis, the spread between the yield to maturity on the short-term German government bond and the short-term Greek bond, a dummy variable equal to 1 if the ECB's quantitative easing is implemented, the level of monthly bond buying, the euro-dollar exchange rate and the volume of press coverage. When it comes to modelling the variance of market participants, I add the conditional volatility of inflation, which is estimated after running a GARCH(1,1) on observed inflation rates, the VIX index, which is an indicator of market stress and the Shannon's entropy of media coverage, which is a measure of uncertainty in information media coverage conveys about the ECB. All variables are collected at the weekly frequency.

Apart from the obvious fundamentals market participants should use to forecast inflation, such as the CPI, the spread between the German and the Greek debt is a proxy for the occurrence of the Greek debt crisis in the summer of 2015. The short-term uncertainty that stemmed from a possibility of default could potentially have driven investors to revise their inflation expectations downwards or to hedge against more inflation scenarii, thereby increasing the variance of market inflation expectations. The Greek crisis was also seen as a test to the Euro Area and a challenge to the ECB's implementation of its monetary policy, making possible the link between a rather localised solvency issue to the aggregate European macroeconomic system.
When forecasting the inflation rate, market participants ought to observe the stance of the central bank's monetary policy. The issue is that, over the whole sample, the ECB's main interest rate is constant and equal to zero. Therefore, one needs to include proxies for the very accommodative ECB's monetary policy. The inclusion of the dummy variable for the implementation of the ECB's asset purchase programme and the inclusion of the level of bond-buying controls for the very likely impacts the ECB's QE may have on market inflation expectations. These impacts could be purely on the state of the inflation option market (for example, impacts on liquidity or on risk premia, or both), and it is also likely that the implementation of QE signalled a future increase in inflation to market participants. The inclusion of the exchange rate is also a proxy for the very accommodative ECB's monetary policy. In particular, the rise of the dollar against the euro is a proxy capturing the deviation between major economic zones' monetary policies. Finally, in other specifications, I include Wu and Zhang's ECB shadow rate.
I include the volume of media coverage in order to capture the volume of events that are newsworthy to central banks. As shown in Picard and al. (2016), volume of media coverage about financial and banking affairs increases in eventful times, therefore it should be able to proxy the occurence of significant events, as well as their impacts on market participants' inflation expectations, over my sample. In particular, it is shown that financial media coverage, which my sample of media stories are mostly drawn from, is more likely to report negative news than positive news. Therefore, I am confident that the volume of ECB media coverage can be a relatively satisfactory proxy for underlying events as well as negative events[27]- such as the Greek crisis, for instance.

[27]In other specifications I replace the raw volume of media coverage by a dummy variable equal to 1 when volume

Finally, I use DTM's time-varying word distributions inside topics to calculate the Shannon's entropy of the ECB media coverage about the ECB's monetary policy. Shannon's entropy is a concept of information theory and is a broad measure of minimum cost to recover some initial message through a noisy channel. Specifically, I calculate Shannon's entropy as

$$H_t(\beta_{t,5}) = -\sum_{v=1}^{V} \beta_{t,k}^{v} Log_2 \beta_{t,k}^{v} = -\sum_{v=1}^{27.104} P(w_v|\beta_{t,5}) Log_2 P(w_v|\beta_{t,5})$$

where $\beta_{t,5}^{v}$ are the time-varying word distributions for topic 5 on the ECB's monetary policy computed in section 3. The higher the entropy of a signal, the higher the minimum amount information an agent needs to recover the initial message and the higher the uncertainty: I therefore take Shannon's entropy as a proxy for the precision of the media signal σ_s^2 from section 2 about the ECB's monetary policy.

Finally, market participants are also able to observe the ECB's communications directly. In order to integrate the ECB's direct communication into the model, it is important to realise that the ECB conveys information to market participants in various forms. Apart from the usual press conference, ECB central bankers engage into regular media activities such as interviews and other speaking engagements. Furthermore, not only does the ECB President directly communicate, but also do other board members.

I choose to model the ECB's direct communication as two variables. Since there are weeks when the ECB does not issue any form of direct communication, the first variable is a dummy equal to 1 when direct communication of any form is issued from the ECB. The second variable is a continuous variable S $\in$ (-1;1), where, in the extreme cases, S takes a value of -1 when the communication is entirely dovish and takes a value of 1 when the communication is entirely hawkish[28].

In order to code the ECB's communications, I use an algorithm called "Wordscore", which is a supervised Bayesian classifier developed by Laver and al. (2003) and that is able to classify incoming documents based on their words' probabilities of appearing in some prior reference texts[29]. I collect the transcripts of all the speeches, interviews and press conferences at the weekly frequency and feed them to Wordscore. Interestingly, most direct ECB communications are found to be neutral- appart from press conferences, that seems to be coded as being regularly dovish. Also, since the stance of monetary policy is mostly dovish over the sample, "hawkish" communications should be more seen as "less dovish" ones, such as communications implying shorter-than-expected unconventional policies or risks to implementing QE.

4 Estimation and results

4.1 Effects on the mean inflation expectation

The series of mean inflation expectation fails to reject the augmented Dickey-Fuller test's null hypothesis[30]. Therefore, one cannot reject the hypothesis that the mean contains a unit root and the series needs to be differenced.

When brought to the data, Hypothesis 1 becomes: the proportion of media coverage dedicated to the ECB's monetary policy affects the level of the mean market inflation expectation when

is above average.

[28] As a reminder, dovish communications imply a more accommodative stance while a hawkish communication imply a tight stance.

[29] For reference texts, I choose the January 22nd 2015 press conference announcing QE as a dovish communication and Mrs Lautenschlager's November 23rd speech on the risks of low rate policies as a hawkish communication. Contrary to topic modelling, supervised algorithms require some amount of human subjectivity prior to classifying new incoming data.

[30] My test statistics under the null is -1.60 and the 10% critical value is -2.60.

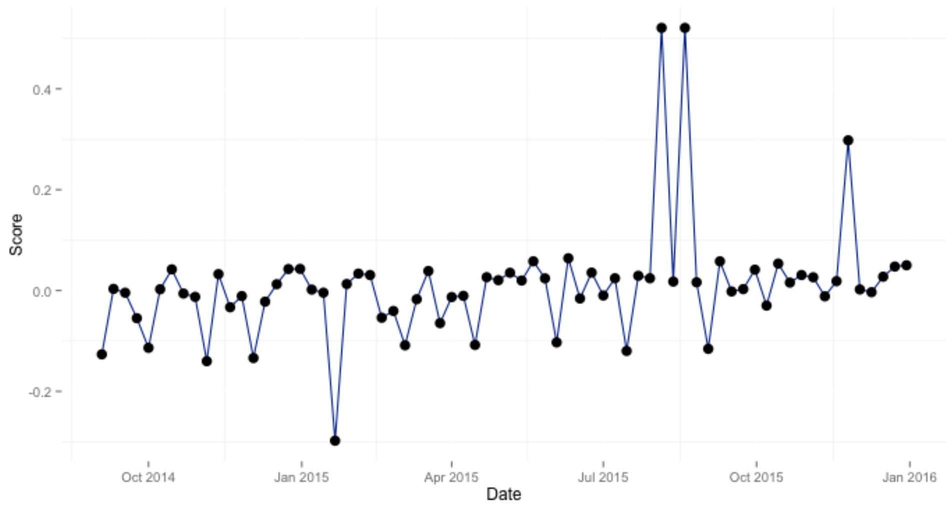

Figure 12: hawkish (1) and dovish (-1) direct ECB communications.

Table 7: Descriptive statistics for other explanatory variables

	N	Min	Max	Mean	Sd
CPI (%)	70	-.6	.4	0.0580	0.265
Shadow rate (%)	70	-3.221738	-1.332135	-2.339	0.428
Monetary refinancing operations (Mn)	70	60.527	165.352	94.85	26.06
QE (Bn)	70	191.24	805.28	415.4	207.3
EUR/USD (1 USD)	70	.77229	.94684	0.875	0.0504
Spread German-Greek debt (ppt)	70	2.053	33.864	12.87	7.661
Volume of media coverage (nb of weekly articles)	70	224	1148	634.5	165.4
VIX	70	11.82	28.03	16.25	3.995
Entropy of media coverage	70	9.646	10.343	10.139	.128
Conditional variance of inflation	70	.0107396	.0119258	0.0110	0.000280
Number of weekly direct communications	70	0	10	2.986	2.216
Score	70	-.2979	.5209	0.00880	0.114

it increases. Therefore, I use OLS to estimate the following regression where all variables are expressed as their first difference:

$$E_t(\pi_{t+1}) = \alpha + \phi E_{t-1}(\pi_t) + \beta_1 CPI_t + \beta_2 SR_t + \beta_3 \Delta QE \tag{9}$$
$$+\beta_4 EUSD_t + \delta_1 DCOM_t + \beta_5 SCORE_t + \beta_6 VOLUME_t + \beta_7 \Delta ECBMP_t + \epsilon_t$$

The results are displayed below and I also display the results from the autoregressive model with an AR(1) for comparison. All relevant variables are found to be significant at least at the 10% level, except for the number of direct communications per week, which is found to be significant in the ARMA model but not in the ARIMA version.

Amongst the traditional information channel of economic data, the CPI and the euro-dollar exchange rate are statistically significant with an expected positive sign: the increase in today's CPI signals an in increase in the future inflation rate. The rise in the USD relative to the EUR also signals an increase in the future inflation rate, which is consistent with a boost to inflation going through exports.

I find that an increase in the volume of media coverage signals a decrease in inflation, which shows that the occurence of ECB-related events throughout the sample decreased market participants' inflation expectations over the one-year horizon.

With regards to conventional and unconventional monetary policy, I find that the regular asset purchases and the shadow rate increase the mean inflation expectation in both models. The effect of QE goes through the traditional monetary policy channel. A rise in the shadow rate could increase inflation expectations as it could signal better-than-expected data through the central bank's reaction function and a rise in future inflation.

When it comes to the ECB's direct communication through the stance of its communication activities, I find that less dovish tones reduces market participants' mean inflation expectations.This result is interesting because the literature on central bank communication- for instance, Hubert (2012) - generally finds that hawkish communications increases inflation expectations through signalling increasing inflation.
But the vast majority of the literature on central bank communication focuses on times when hawkish communications implied dangerously rising inflation through communicating about how necessary it is to raise interest rates. For instance, works like Jansen and De Haan (2007) study the future monetary policy path implied by hawkish statements in counting the occurrences of the word "vigilance" in official ECB communications. "Vigilance" was used by Jean-Claude Trichet in order to signal that inflation had reached a level such that the ECB had to raise interest rates in order to cope with its medium-term target of 2%. Of course such times are far past and my sample covers a time of opposite risks, namely stagnating low inflation rates, decreasing inflation expectations and material deflation risks. In my sample, most hawkish communications are therefore only less dovish ones and they never imply underlying accelerating inflation and future interest rates hikes. In particular, examples of "hawkish" communications are communications on the risks of purchasing government bonds. Therefore, in my case, hawkish communications do not signal increasing inflation data but rather, possible impeachment to more unconventional measures.
Therefore, my sample's "less dovish" communications do not imply signals about the state of the economy through the central bank's reaction function and are interpreted through the transmission mechanism of monetary policy: the attached negative sign on "hawkish communications" is consistent with the fact that accommodative monetary policy should induce higher inflation in the future, and conversly, that a less accommodative stance could imply reduced future inflation.

In both models, the coefficient attached to the (changes in) proportion of media coverage about the ECB that is dedicated to the ECB's monetary policy is found to be statistically significant and and positive, which is consistent with theory and hypothesis 1: the mean inflation expectation is affected by positive changes in the proportion of media coverage dedicated to the ECB's monetary policy. Media coverage about the ECB's monetary policy is therefore found to be informative for market participants to forecast inflation over the one-year horizon- even after controlling for direct communication, economic data and the occurence of ECB-related events throughout the sample (via volume of ECB media coverage).
The nature of this information is hard to tell from the model and could be either of a distorting nature or of an amplifying nature with regards to data and the ECB's direct communication. The positive sign on the media signal may imply that the media signal is mostly dovish in its stance, if market participants interpreted the media signal as being about the monetary transmission mechanism and the effects of monetary policy. In this case, the media signal could be thought as being significant to the formation of market participants' inflation expectations by being an amplification device for the initial ECB's direct communication. But if market participants mostly interpreted media coverage about the ECB's monetary policy as signals about improving economic data through the central bank's reaction function, the same positive sign attached to the media signal could also imply that the signal is "less dovish" and media coverage about the ECB's monetary policy could be distortive of the ECB's initial direct communication.
The present model does not allow to disentangle the two channels but, given the timeframe of the sample and our interpretation of the effects of the ECB's direct communication, the hypothesis that media coverage may be amplifying the ECB's initial stance seems more realistic.

Table 8: Test of hypothesis 1 with respect to mean expectations. Autoregressive and integrated models.

	H1	(1) ARMA(1,0,0)		(2) ARIMA(0,1,0)	
Dep. var: mean					
Lagged dependent variable		0.635***	(0.0881)	-	-
CPI		0.332***	(0.0876)	0.379***	(0.101)
Shadow rate		0.231***	(0.0831)	0.330**	(0.154)
Δ QE		0.00671*	(0.00393)	0.0123**	(0.00564)
EUR/USD		2.235***	(0.614)	1.332*	(0.797)
Volume media		-0.000231***	(0.0000816)	-0.000138*	(0.0000788)
Direct communication		0.0411*	(0.0210)	-0.00647	(0.00472)
Score		-0.0399***	(0.00830)	-0.0286***	(0.00730)
Δ ECB MonPol	Significant	0.676**	(0.326)	0.398*	(0.211)
Constant		-1.235***	(0.376)	-0.000807	(0.0139)
N		68		67	
adj. R^2		0.947			

All dependent variables are expressed as first differences in the case of the integrated model.
In the autoregressive model, the variable for direct communication is a dummy equal to 1 when the ECB engages into any forms of direct communication and 0 otherwise. In the integrated model, the variable for direct communication is the number of direct communication activities the ECB engages into within a given week. Robust standard errors in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

4.2 Effects on the variance

The variance of inflation expectations also fails to reject the null hypothesis of an augmented Dickey-Fuller test that it contains a unit root and the series of variances is therefore differenced. In the context of my data, hypothesis 2 becomes: the changes in proportion of ECB media coverage dedicated to the ECB's monetary policy affect (and reduce) the variance of market participants' inflation expectations. Since the entropy of media coverage is high when uncertainty is high, hypothesis 3 is: the entropy of media coverage about the ECB's monetary policy is positively correlated to the variance of market participants' inflation expectations.
In order to account for the high volatility of the variance in inflation expectations, the variance is also modelled using an ARCH(1) of the form[31]:

$$\begin{aligned} VAR_t(\pi_{t+1}) = \alpha + \beta_1 VOL + \beta_2 CPI_t + \beta_3 SR_t + \beta_4 \Delta QE \\ +\beta_5 EUSD_t + \beta_6 DCOM + \beta_7 EURUSD_t + \beta_8 VIX_t + \beta_9 SCORE_t \\ +\beta_{10} VOLUME_t + \beta_{11} \Delta ECBMP_t + \beta_{12} ENTROPY + u_t \end{aligned} \tag{10}$$

where each variable is expressed as its first difference and $u_t = \sqrt{(h_t)} \cdot v_t$, $h_t = a + bu_{t-1}^2$ and $v_t \sim \mathcal{N}(0,1)$. The results from this ARIMA(0,1,0)-ARCH(1) are displayed below as well as those from the autoregressive ARMA-ARCH model, alongside the expected signs of the variables of interest implied from hypotheses 2 and 3.

With regards to economic data, an increase in CPI and in the EUR/USD exchange rate is found to decrease the variance of inflation expectations. The estimated variance of inflation expectation may stem from uncertainty or disagreement amongst market participants towards the future inflation rate. Therefore, if we combine these results with the result on the mean

[31]One rejects the null of no ARCH effect in the residuals of the regression in differences.

inflation expectations, we have that an increase in the CPI and in the exchange rate had investors increase their inflation expectation with more certainty or less disagreement. Unlike the CPI, an increase in the VIX increases the variance of inflation expectations, which is consistent with the VIX' being an indicator of market stress, thereby affecting market uncertainty.

An increase in the shadow rate is found to increase the variance of inflation expectations. If we combine this result with the result that the shadow rate also increases the mean expectation through signalling improving data, we may say that an increase in the shadow rate has market participants increase their inflation expectations but with more uncertainty around the mean.
While the regular bond purchases are found not to explain the variance of market participants' inflation expectations, the stance in the ECB's direct communication does. A less dovish stance in the ECB's direct communication increases the variance of market participants' expectations. Again, this is to be interpreted alongside the result that less dovish communications decrease the mean inflation expectation. In a dovish policy environment and regular dovish stances, punctual more hawkish communications may input uncertainty about the future inflation rate through a perceived tightening via the perceived functioning of monetary policy.

Hypothesis 2 is not confirmed by the data: an increase in the proportion of the ECB media coverage dedicated to the ECB's monetary policy does affect the variance of market participants' inflation expectations but it increases it. The reason why the theory predicts such an hypothesis lies in the way it assumes the media signal is informative about the future inflation rate. For instance, in Carroll (2003) the media is seen as a mere recipient for professional forecasters' expectations to be spread to consumers. In this context, the more media coverage the less dispersed the expectations because professional expectations are conveyed without noise. But in reality, media are not noise-free information channels and they do not work in simply distributing data to the outside world.
Two pieces of theory may explain why I find that more media coverage about the ECB's monetary policy increases the variance of market participants' inflation expectations. Because of the man-bites-dog effects, extreme events have more probabilities to be picked up by media than common events. Therefore, the perceived probability that the future inflation rate will be uncommon when the ECB's monetary policy is reported in the media is higher under man-bites-dog than under a uniform equal treatment of events. From the agent's perspective, observing media coverage about the ECB's monetary policy may increase perception of tail-probability events and increase market participants' uncertainty/disagreement about the future inflation rate[32].
Morris and Shin (2012) also provides some intuition for why an increase in media coverage about the ECB's monetary policy may increase the variance of inflation expectations. When agents possess high-quality private information about a variable, the provision of a public signal about it may lead to overreaction to noise. In the context of the ECB's monetary policy, it is possible that market participants perceive too much uncertainty from particular media narratives, which could lead to the increase of the observed variance.

Hypothesis 3 is supported by the data: the uncertainty of media coverage about the ECB's monetary policy is found to increase the variance of market participants' inflation expectations. The entropy of the ECB media coverage about the ECB's monetary policy captures how informative media coverage about the ECB's monetary policy is. "Informative" refers to information theory, where it relates to the minimum cost (or bits, information units) an agent incurs in order to recover some initial message out of a noisy channel. Therefore, the higher the media entropy, the less informative media coverage about the ECB's monetary policy is. In the context of topic modelling, the entropy captures how "useful" the time-varying word distributions are at each point in time in conveying information about the ECB's monetary policy.

The entropy of media coverage seems to oscillate around some mean value with clear lower values in the beginning and at the end of the sample as well as during the summer of 2015. In the context of the model presented in section 2, the reason why an decrease in entropy may decrease the variance of inflation expectations is because after observing the media signal, market

[32]The effect of an-bites-dog on expectations is described in Perrera das Neves 2016b.

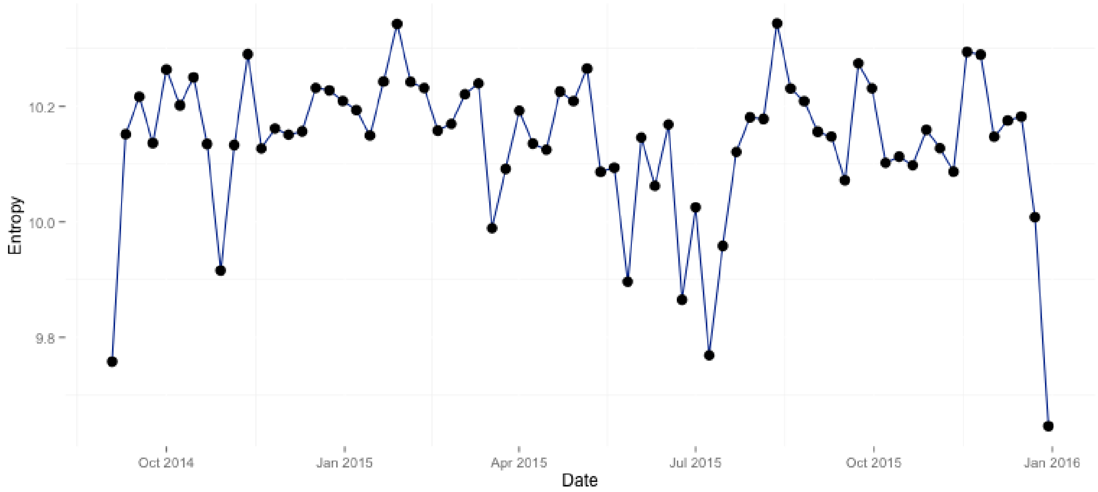

Figure 13: Shannon's entropy of media coverage about the ECB's monetary policy. Lower entropy means more clarity with regard to the underlying message.

participants update their prior belief about the future inflation rate. Their posterior belief is a weighted average about their prior information and the media signal - and the variance of their posterior is an decreasing function of the precision of the media signal.
When the media signal is precise (entropy is low), the media delivers peaked distributions of words that are informative about the most likely state of future inflation agents are in and market participants' uncertainty/disagreement about the future inflation rate decreases.

5 Robustness checks

5.1 Time series aspects

In this section, I show why the mean expectations and the variance were modelled using an ARIMA(0,1,0) instead of an ARMA(1,0).
After differencing the series in order to have stationary data, I plot the autocorrelations and partial autocorrelations functions of the mean inflation expectations and the variances of expectations. It is clear that there is no persistence in either series' autocorrelation functions, therefore, no autoregressive term is required in order to model the two series. Looking at the partial autocorrelation functions, the same observations apply and no moving average term in required either. Cross-checking the graphs, the ACP and PACF of the mean and the variance of inflation expectations are consistent with an ARIMA(0,1,0) model.

The residuals appear to be nicely distributed around zero, although the residuals on the regression on the variance display some clear outliers. Both series do not display any serial correlation.

5.2 Model specification

In this section, I run several specification of the same two core regressions for the mean inflation expectation and the variance of expectations as presented above. First, I iterate between AR-ARCH models in level for the sake of comparability as most of the literature on central bank communication presents results in levels, and first-differenced ARIMA models. Secondly, I test whether the inclusion of key control variables change the significance and the magnitude of my estimates of interest. Thirdly, I test whether adding new control variables changes the magnitude and the significance of my estimates of interests. Fourthly, I test whether using the changes in

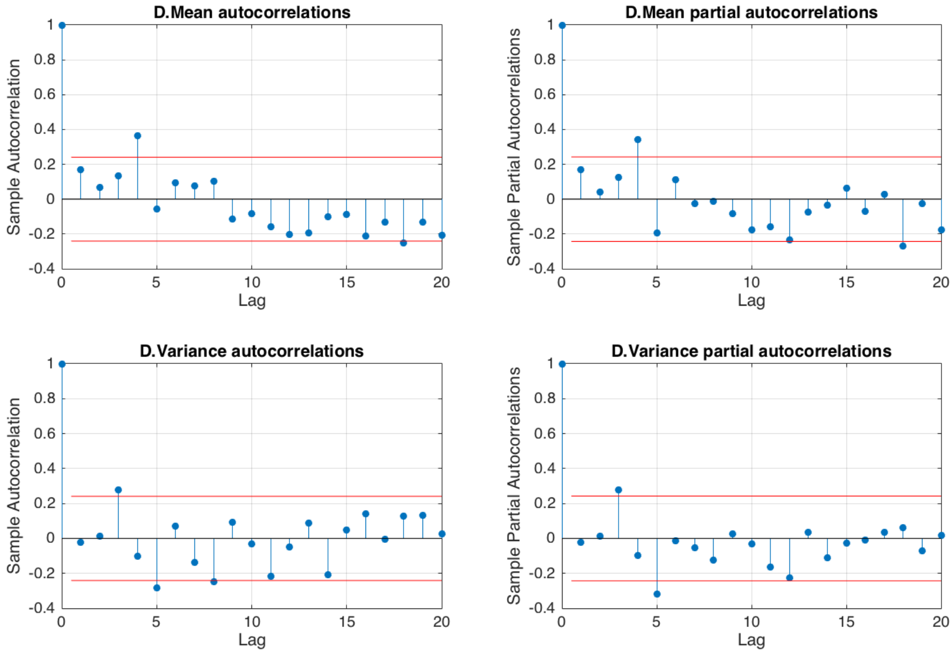

Figure 14: Autocorrelation and partial autocorrelations of the first differences of the mean/variance of expectations.

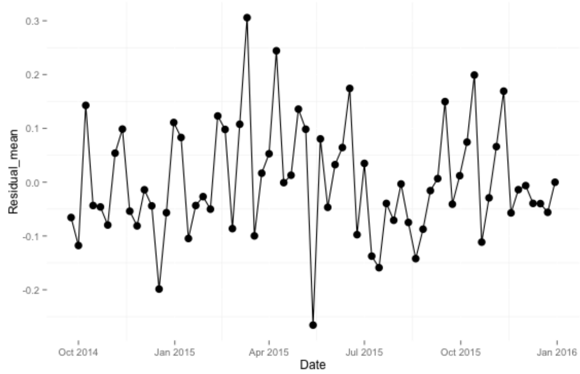

Figure 15: Residuals from the regression on the mean (ARIMA)

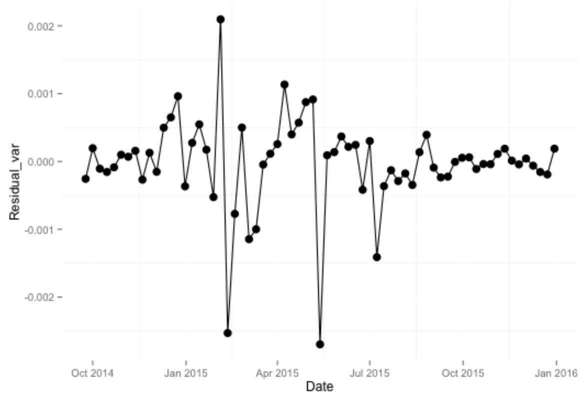

Figure 16: Residuals from the regression on the variance (ARIMA)

Table 9: Test of hypotheses 2 and 3 with respect to the variance of expectations. Autoregressive and integrated models.

		(1) ARMA-ARCH		(2) ARIMA-ARCH	
	H2/H3				
Dep.var: Variance					
Conditional volatility		2.362***	(0.0840)	1.143***	(0.273)
CPI		-0.00110***	(0.0000587)	-0.00113***	(0.000146)
VIX		0.0000225***	(0.00000870)	0.00000904*	(0.00000549)
EUR/USD		0.00148	(0.00109)	-0.00131	(0.00107)
Shadow rate		0.000284*	(0.000151)	0.000456***	(0.000154)
Δ QE		0.00000280	(0.00000401)	-0.00000695	(0.00000762)
Volume media		0.000000121**	(4.91e-08)	0.000000271**	(0.000000106)
Score		-0.00000641	(0.0000137)	0.0000344***	(0.0000110)
Δ ECB MonPol	(-)	0.00125***	(0.000297)	0.000776**	(0.000370)
Entropy	(+)	0.000394***	(0.000117)	0.000521***	(0.000163)
Constant		-0.0225***	(0.00171)	0.0000257	(0.0000163)
ARMA					
AR(1)		0.946***	(0.0236)	-	-
ARCH					
ARCH(1)		3.053***	(0.956)	2.943***	(0.937)
Constant		3.66e-09	(3.02e-09)	2.15e-09	(3.35e-09)
N		68		67	

Robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

proportions of some wording inside the ECB's monetary policy topic instead of the changes in proportion of the ECB's monetary policy topic itself changes the significance, the magnitude and the interpretation of the effect of the media signal.

The estimates for the effect of the media signal on the mean inflation expectation seems to be constant throughout all the specifications. In particular, its sign is always positive, consistently with hypothesis 1. The magnitude and the statistical significance does not seem to be affected by the inclusion of new regressors such as the spread between the short-term German and Greek debt or the industrial production. Actually, the estimate hardly changes when no control variable is included. The inclusion of the proportion of the wording "buy government bond" inside the topic about the ECB's monetary policy instead of the topic proportion in the ECB media coverage itself does not change the interpretation of the effect of the media signal on the mean inflation expectation and hypothesis 1 is still validated by the data.
On a related note, the estimates attached to direct communications' stance remain unaffected by the various specifications of the model. The sign on the total volume of media coverage about the ECB remains negative throughout the specifications but is found to be insignificant when varying control variables.
All in all, I find that the estimation of the role of the media signal (as well as the role of direct communication, to this regard) is robust to varying the model's specification.

The estimates for the effect of the media signal on the variance of market participants' inflation expectations is also rather robust to various model specifications, although some specifications reject hypothesis 3 and do not confirm the core regression's results about hypothesis 2.

When it comes to the effect of the proportions of the topic about the ECB's monetary policy in the total ECB coverage on the variance of inflation expectations, both models in level deliver coherent estimates with regard to the core regression presented above. That is, they both reject

hypothesis 2 that the media signal decreases the variance of inflation expectations and deliver positive coefficients attached to the media signal. This result remains unchanged by the inclusion of the proportion of wording "expand balance sheet" inside the topic about the ECB's monetary policy instead of the proportion of the topic itself. I interpret this result in the same way as I did earlier (man-bites-dog and/or overreaction to public signal).
The estimates show less robustness when considered in the first-differenced model. The effect of the change in the proportion of the topic about the ECB's monetary policy is found to be insignificant in the context of a first-differenced model with additional regressors. Interestingly, the inclusion of changes in the wording "buy government bonds" inside the topic about the ECB's monetary policy in the context of the first-differenced model while controlling for key monetary policy variables, market stress and the starting week of the ECB's asset purchase programme, yields a negative coefficient on the effect of the media signal, thereby confirming hypothesis 2.

Hypothesis 3 is confirmed in both models in level and is robust to the inclusion of new regressors in the first-differenced model. But entropy is found to be insignificant when including changes in the wording "buy government bonds" instead of changes in the underlying topic proportions. One explanation could be that the proportions of the wording "buying government bonds" inside the topic about the ECB monetary policy also explains changes in entropy. That would imply that when the media covers bond purchases to describe the ECB's monetary policy, it is less costly for market participants to observe the informational content the media signal about the ECB's monetary policy conveys. This is mildly confirmed by simple regressions.

All in all, I find that the role of the media signal in the variance of market participants' inflation expectations is mostly robust although I acknowledge that some effects are more dependent on the specification of the model than the coefficients on the mean expectations. In particular, my estimates seem robust throughout the models in level, which is important in the context of the literature on central bank communication. Indeed, most of the studies of the effects of bank communication are conducted in levels. The estimates in first-differenced models are more dependent on the inclusion of other regressors but the key difference to the core regressions occur with including the media signal's particular wording instead of the topic itself, which suggests that the sign of the media signal may also depend on the content of the signal - which is not at odd with theory.

The appendix displays the estimates of the effects of the media signal on market participants' implied probabilities of deflation over the one-year horizon. The media signal about the ECB's monetary policy is found to decrease the implied probability of deflation while the overall volume of media coverage is found to increase it, which is consistent with previous findings. Results on the effect of the media signal about the ECB's monetary policy are unchanged whether I include changes in the topic proportions or changes in the wording inside the topic. Hawkish stance in the ECB's direct communication increases market participants' probability of deflation, which is also consistent with findings on the effects of such tones on the mean inflation expectations as well as on the variance.

Lastly, some limitations should be noted. First of all, due to limitations in gathering media data, the sample account for 70 observations which, in the context of the number of regressors, may affect the quality of inference. One improvement could come from estimating the same models via bayesian regressions because providing a prior distribution for each variable - even a flat one - should help decrease standard errors in the context of a low degree of freedom.
Another limitation is that of the possible endogeneity of the media variables. Even if I control for major monetary policy moves through data and for key economic events throughout the sample using the volume of ECB media coverage, it is still possible that my media variables could be correlated to some unobserved events. Hubert (2016, 2017) propose one way of addressing this issue in drawing from Romer and Romer (2004)'s identification strategy. Exogenous shocks are derived from series of possibly endogenous regressors in computing the residuals from regressions of endogenous variables on a list of exogenous regressors. I keep this approach for further work.

Table 10: Robustness checks for the mean inflation expectation.

begintabularl*4cc	(1)		(2)		(3)		(4)	
	AR(1)		No AR		ARIMA		ARIMA	
	With controls		No Control		Wording		Other controls	
Dep.Var=Mean								
Lagged mean	0.693***	(0.0804)						
CPI	0.00341***	(0.00113)			0.431***	0.0986)	0.370***	(0.102)
Shadow rate	0.00162	(0.000992)			0.393**	(0.175)	0.337**	(0.157)
QE	-0.000000948	(0.00000144)			0.0152**	(0.00627)	0.0126*	(0.00672)
EUR/USD	0.0256***	(0.00652)			1.361	(0.868)	1.254	(0.820)
Direct Communication	0.0000228	(0.0000561)	-.0030	(0.004)	-0.00652	(0.00499)	-0.00442	(0.00496)
Volume	-0.00000241**	(0.000000950)	-.0000919	(0.00007)	-0.000148**	(0.0000729)	-0.000111	(0.0000790)
Score	-0.000283***	(0.0000574)	-.0243863***	(0.007)	-0.0231***	(0.00713)	-0.0264***	(0.00764)
Δ ECB MonPol	0.00593*	(0.00304)	.548009*	(0.30)			0.411*	(0.215)
Δ "buy government bonds"					123.6*	(75.24)		
Industrial prod							-0.00276	(0.0170)
Spread							-0.00464	(0.00390)
Constant	-0.0153***	(0.00424)	-0.004	(0.014)	0.00104	(0.0141)	0.000163	(0.0139)
sigma								
Constant					0.104***	(0.0106)	0.103***	(0.0100)
N	68		68		67		67	
adj. R^2	0.943							

Robust standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Table 11: Robustness checks for the variance of inflation expectations.

	(1) AR-ARCH No controls		(2) AR-ARCH Wording		(3) ARIMA Other controls		(4) ARIMA Wording	
Dep.var=Variance								
Volume	2.22e-07***	(1.26e-08)	-1.44e-07***	(3.58e-08)	2.35e-07***	(8.72e-08)	6.80e-07	(4.09e-07)
Score	2.85e-05***	(3.83e-06)	1.93e-05*	(1.06e-05)	2.74e-05**	(1.20e-05)	1.02e-04***	(3.36e-05)
Δ ECB MonPol	4.10e-04***	(0.000109)			5.04e-04	(3.64e-04)		
Entropy	8.15e-04***	(0.000237)	8.84e-04***	(0.000130)	2.33e-04**	(9.42e-05)	3.33e-04	(0.000464)
Conditional volatility			1.486***	(0.122)	1.225***	(0.230)	-0.350	(0.818)
CPI			-0.00105***	(0.0000662)	-0.00125***	(0.000154)	-0.00135	(0.000868)
Shadow rate			-0.0000438	(0.0000420)				
Δ QE			-0.0000245***	(0.00000344)	-0.0000145	(0.0000145)		
Δ "expand balance sheet"			0.0450**	(0.0225)				
Spread					0.0000185***	(0.00000529)		
EUR/USD					-0.000439	(0.00155)	-0.00806*	(0.00429)
MRO					0.00000246**	(0.00000101)	0.00000651	(0.00000705)
VIX							0.0000301*	(0.0000169)
Start QE=1							-0.000769***	(0.000143)
Δ "buy government bonds"							-1.383***	(0.497)
Constant	8.79e-03***	(0.000795)	-0.0110***	(0.00137)	-0.0000160	(0.0000349)	0.0000201	(0.0000788)
ARMA								
AR(1)	0.938***	(0.0116)	0.553***	(0.0286)				
ARCH								
ARCH(1)	5.502***	(1.517)	3.425***	(0.855)	2.463***	(0.571)		
Constant	1.78e-09	(1.26e-09)	7.32e-10	(1.62e-09)	7.62e-09*	(4.35e-09)		
N	69		68		67		68	

Robust standard errors in parentheses

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

6 Conclusion

The ECB attracts a lot of media coverage as it is at the foreground of today's policy-making. Appart from building the ECB's reputation and image in the public, media convey most of the ECB's communications. But how informative is media coverage about the ECB for market participants to form inflation expectations, after they observe economic data and the ECB's communications directly ? In this paper, I answer this research question considering a framework in which market participants could partly "learn" from media in order to update their prior belief about the future inflation rate, even at near-zero information costs in the presence of finite information processing capacity in the spirit of the literature on rational inattention. The main finding of this paper is that media coverage is found to bring some useful information about the future inflation rate, alongside other traditional information channels.

In particular, the media signal about the ECB's monetary policy is found to increase market participants' inflation expectations and variance of inflation expectations. While the former result is consistent with bayesian learning, the latter result poses a challenge to the kind of information on the future inflation rate media convey to market participants. For example, in the presence of man-bites-dog effect, media coverage of an event increases its probability to turn out to be uncommon, from the agent perspective. This effect is also consistent with Morris and Shin (2012) who show that otherwise well-privately-informed agents may overreact to public signals that give too much weight on noise.
The final effect of the media signal may depend on its content as I find that changes in some wordings inside media's topic about the ECB's monetary policy reduce market participants' implied probabilities of deflation and the variance of their inflation expectations, thereby acting much like a "focal point" for market participants' beliefs about the future inflation rate. I finally find that the uncertainty of the ECB media coverage about the ECB's monetary policy increases the variance of inflation expectations, which is consistent with bayesian learning.

This work sheds light on the importance of media in the monetary policy transmission mechanism, in affecting market participants' inflation expectations. In a context in which some constraint on information processing is binding, media coverage about the ECB's monetary policy may produce signals about the most likely state of future inflation to market participants. How clear ECB central bankers and the ECB's initial communications are may also decide whether media will act in focusing market belief or in blurring it.

Appendix

Main topics over time

Table 12: Some instances of $\beta_{0,t}$ "UK Politics"

Sep 29-Oct 5 14'	12-18 Jan 15'	14-20 Sep 15'	Nov 30-Dec 6 15'
prime minister	prime minister	european union	prime minister
mr cameron	member states	prime minister	member states
lord hill	european union	refugee crisis	mr cameron
david cameron	mr cameron	border controls	european union
european union	european commission	asylum seekers	david cameron
conservative party	general election	member states	angela merkel
european parliament	foreign policy	mr corbyn	single market
mr osborne	european commissioner	mr cameron	downing street
party conference	eu member	migrant crisis	inwork benefits
general election	eu membership	border serbia	schengen area

Table 13: Some instances of $\beta_{1,t}$ "Other central banks"

Sep 29-Oct 5 14'	12-18 Jan 15'	14-20 Sep 15'	Nov 30-Dec 6 15'
central bank	central bank	federal reserve	central bank
interest rates	united kingdom	interest rate	interest rates
hong kong	oil prices	emerging markets	federal reserve
federal reserve	interest rates	raise rates	janet yellen
central banks	central banks	rate hike	rate hike
monetary policy	reserve bank	the fed	first time
reserve bank	interest rate	janet yellen	us dollar
wall street	federal reserve	financial markets	us economy
us dollar	new york	us economy	raise interest
fed officials	united states	rate increase	raise interest rates

Table 14: Some instances of $\beta_{2,t}$ "International economy"

Sep 29-Oct 5 14'	12-18 Jan 15'	14-20 Sep 15'	Nov 30-Dec 6 15'
economic growth	economic growth	international monetary	international monetary
public finances	domestic product	interest rates	interest rates
international monetary	interest rates	emerging markets	exchange rate
gross domestic	gross domestic	domestic product	gross domestic product
budget deficit	christine lagarde	gross domestic product	economic growth
domestic product	financial crisis	economic cooperation	budget deficit
financial crisis	public sector	global economy	global economy
international monetary fund	global economy	economic recovery	public finances
global economy	united states	cooperation development	oil prices
spending cuts	france italy	emerging economies	fiscal policy

Probability of deflation over the 1-year horizon

Table 15: Some instances of $\beta_{5,t}$ "ECB's monetary policy"

Sep 29-Oct 5 14'	12-18 Jan 15'	14-20 Sep 15'	Nov 30-Dec 6 15'
european central bank	european central bank	european central bank	european central bank
mario draghi	quantitative easing	euro area	deposit rate
assetbacked securities	government bonds	quantitative easing	quantitative easing
balance sheet	monetary policy	euro zone	mario draghi
covered bonds	european court justice	mario draghi	cut deposit rate
quantitative easing	swiss national bank	interest rates	eurozone economy
governing council	euro zone	qe programme	government bonds
president mario draghi	exchange rate	government bonds	ecb will
government bonds	interest rates	covered bonds	rate cut
interest rates	swiss franc	nonperforming loans	negative territory

Table 16: Some instances of $\beta_{8,t}$ "European politics"

Sep 29-Oct 5 14'	12-18 Jan 15'	14-20 Sep 15'	Nov 30-Dec 6 15'
european union	mr tsipras	new democracy party	deposit insurance
finance minister	prime minister	mr tsipras	capital controls
european commission	alexis tsipras	syriza party	finance ministers
prime minister	syriza party	european union	austerity measures
angela merkel	general election	finance minister	european central bank
mr hollande	european central bank	debt relief	socialist party
national front	greek exit	new government	spending cuts
mr moscovici	greek government	capital controls	international monetary fu
french finance minister	international monetary fund	bailout deal	banking union
mr renzi	greek banks	greeces creditors	tax increases

Table 17: ARIMA(0,1,0) and AR(1) models for market participants' implied probability of deflation over 1 year.

	(1)		(2)		(3)		(4)	
	ARIMA		ARIMA		AR		AR	
	Topic		Wording		Topic		Wording	
Dep.var=Pr. deflation								
CPI	-0.145**	(0.0590)	-0.207***	(0.0532)	-0.188***	(0.0376)	-0.186***	(0.0415)
Shadow rate	-0.0999*	(0.0580)	-0.118*	(0.0604)	-0.0836***	(0.0304)	-0.0934***	(0.0311)
Δ QE	-0.00248	(0.00175)	-0.00361*	(0.00186)	-0.00170	(0.00160)	-0.00215	(0.00162)
EUR/USD	-0.483*	(0.261)	-0.499*	(0.283)	-0.764***	(0.224)	-0.794***	(0.219)
Volume	0.0000438	(0.0000327)	0.0000480**	(0.0000234)	0.0000975***	(0.0000318)	0.0000881***	(0.0000291)
Score	0.00881***	(0.00260)	0.0104***	(0.00243)	0.0172***	(0.00299)	0.0141***	(0.00261)
Δ ECB MonPol	-0.151*	(0.0794)			-0.211*	(0.113)		
Start QE=1	-0.178***	(0.0206)	-0.177***	(0.0201)	-0.115***	(0.0211)	-0.131***	(0.0169)
Δ "buy government bond"			-73.01***	(14.83)			-54.66*	(28.05)
Lagged dependent variable					0.562***	(0.0889)	0.533***	(0.0980)
Constant	0.00406	(0.00545)	0.00363	(0.00488)	0.578***	(0.168)	0.601***	(0.166)
N	66		66		67		67	
adj. R^2	0.466		0.570		0.948		0.950	

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

The shifted SABR/Black model

The SABR is a stochastic model of volatility used to model any forward rate (in my case, the inflation swap rate which is the ATM strike of inflation cap options). In particular, it aims at estimating the implied volatility curve called the volatility smile.

The SABR model of Hagan et al. is described by the following 3 equations

$$\begin{aligned} df_t &= \alpha_t f_t^\beta dW_t^1 \\ d\alpha_t &= v\alpha_t dW_t^2 \\ E\left[dW_t^1 dW_t^2\right] &= \rho dt \end{aligned} \tag{11}$$

with initial values f_0 and $\alpha = \alpha_0$. and where f_t is the forward rate, α_t is the volatility, and W_t^1 and W_t^2 are correlated Brownian motions, with correlation ρ. The parameters are

- α the initial variance
- v the volatility of variance
- β the exponent for the forward rate
- ρ the correlation between the Brownian motions.

The prices of European call options in the SABR model are given by Black's model. For a current forward rate f, strike K, and implied volatility σ_B the price of a European call option with maturity T is

$$C_B(f, K, \sigma_B, T) = e^{-rT}\left[fN(d_1) - KN(d_2)\right] \tag{12}$$

with

$$d_{1,2} = \frac{\ln f/K \pm \frac{1}{2}\sigma_B^2 T}{\sigma_B\sqrt{T}}$$

and analogously for a European put. The volatility parameter σ_B is provided by the SABR model. With estimates of α, β, v, and ρ, the implied volatility is

$$\begin{aligned} \sigma_B(K,f) &= \frac{\alpha\left\{1 + \left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta v\alpha}{(FK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}v^2\right]T\right\}}{(fK)^{(1-\beta)/2}\left[1 + \frac{(1-\beta)^2}{24}\ln^2\frac{f}{K} + \frac{(1-\beta)^4}{1920}\ln^4\frac{f}{K}\right]} \\ &\quad \times \frac{z}{\chi(z)} \\ z &= \frac{v}{\alpha}(fK)^{(1-\beta)/2}\ln\frac{f}{K} \\ \chi(z) &= \ln\left[\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right]. \end{aligned} \tag{13}$$

Once the parameters α, β, ρ, and v are estimated, the implied volatility σ_B is a function only of the forward price f and the strike K. It is therefore possible to use the SABR's estimated coefficients to extrapolate the volatility smile at unobserved strikes and to interpolate the volatility smile between observed strikes.

The β parameter may be chosen amongst values $=$ (0,0.5,1) and does not matter much for the final result. There are two methods for estimating α, ρ, and v, and this paper chooses to estimate all parameters directly. It is also possible to estimate ρ and v directly, and infer α from ρ, v, and the at-the-money volatility, σ_{ATM}. The choice of method does not change the final result.

From equation (13), the at-the-money volatility σ_{ATM} is obtained by setting $f = K$ in equation (13), which produces

$$\sigma_{ATM} = \sigma_B(f,f) = \frac{\alpha\left\{1 + \left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{f^{2-2\beta}} + \frac{1}{4}\frac{\rho\beta v\alpha}{f^{1-\beta}} + \frac{2-3\rho^2}{24}v^2\right]T\right\}}{f^{1-\beta}}. \tag{14}$$

In practice, the choice of β has little effect on the resulting shape of the volatility curve produced by the SABR model, so the choice of β is not crucial. I use β=0.5.

Once $\hat{\beta}$ is set, it remains to estimate α, ρ, and v. This can be accomplished by minimising the errors between the model and market volatilities $\{\sigma_i^{\text{mkt}}\}$ with identical maturity T.

$$(\hat{\alpha}, \hat{\rho}, \hat{v}) = \underset{\alpha,\rho,v}{\arg\min} \sum_i \left\{\sigma_i^{\text{mkt}} - \sigma_B(f_i, K_i; \alpha, \rho, v)\right\}^2. \tag{15}$$

It is therefore possible to use α, β, ρ, v in equation (13) to obtain a continuous vector of σ_B. σ_B can then be used into Black's formula (12) to get the continuous call price function.

Both the SABR and the Black model assume that the underlying forward rate is distributed on positive values only, making it possible to apply the logarithm function in equations (12) and (13). In the current context of close-to-zero interest rates, and low inflation rates, strikes of inflation caps are generally distributed around 0 and negative strikes are traded. Furthermore, the forward rate - the inflation swap rate- may become negative from time to time, as is the case in my sample. The shifted SABR and Black models offer a short-term "patch" to this issue by further assuming :

$$\begin{aligned} df_t &= \alpha_t (f_t + s)^\beta dW_t^1 \quad (16) \\ d\alpha_t &= v\alpha_t dW_t^2 \end{aligned}$$

$$(17)$$

essentially assuming that the forward rate follows a process that is shifted by some positive value s>0. The shift enters (13) and (14) in a simple additive manner to the forward and to the strike price, making it possible to integrate negative strikes and forward rates into the SABR formulas.

Market inflation expectations and SPF

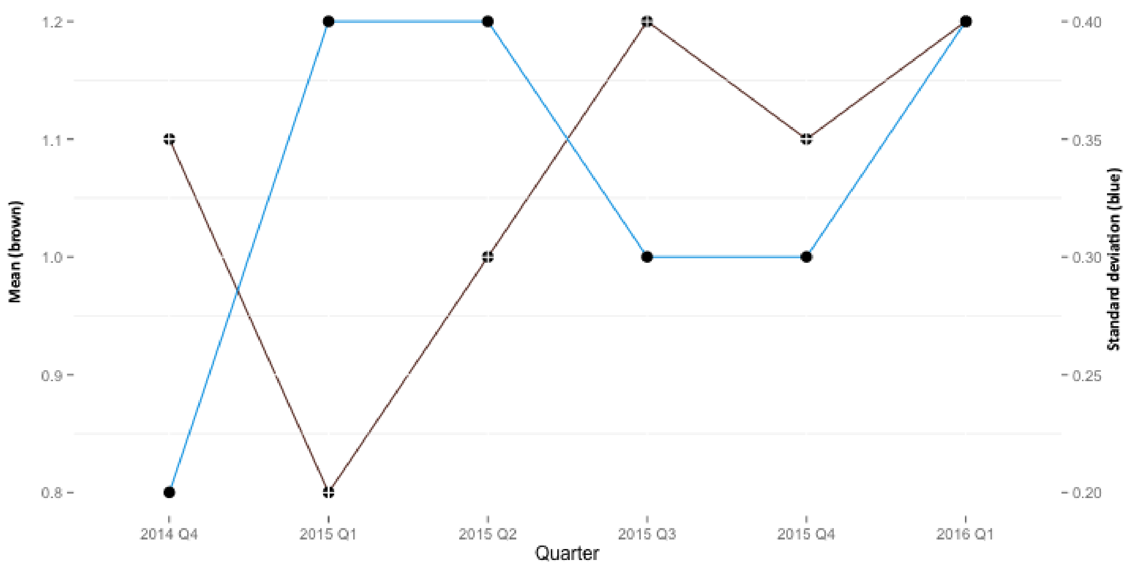

Figure 17: Mean SFP 1y inflation expectation (brown line) and standard deviation (blue line) over the sample.

The European survey of professional forecasters provide quarterly point estimates for professional forecasters' inflation expectations over the one-year horizon as well as their standard

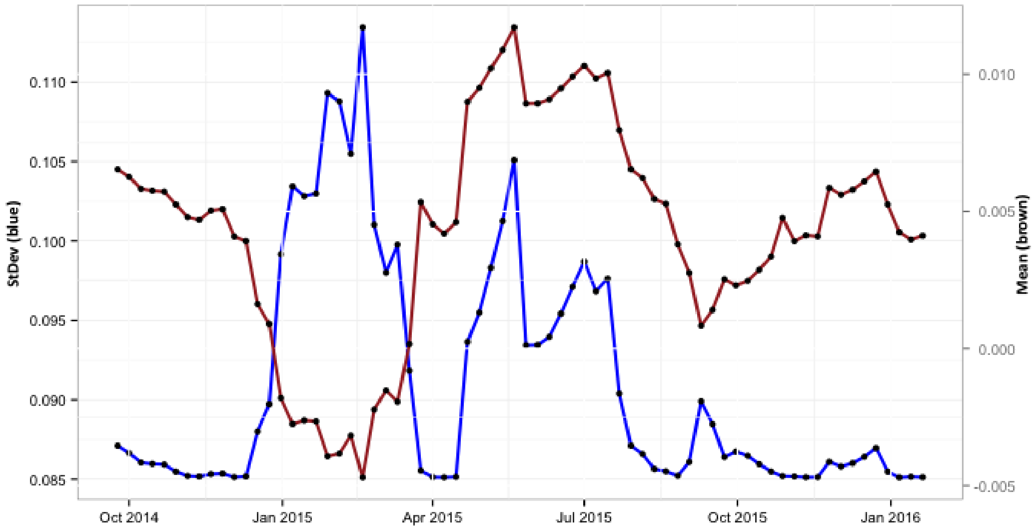

Figure 18: Mean market 1y inflation expectation (brown line) and standard deviation (blue line) over the sample.

deviation. SPF's expectations display much less variations both in mean and standard deviation but exhibit very similar patterns to those in market participants'. Interestingly, unlike market participants, professional forecasters never formed deflation expectations over the sample.

References

[1] Altavilla et al. *Bond Returns and Market Expectations* journal of Financial Econometrics 12 (4): 708-729, 2014.

[2] Andersen and Wagener *Extracting Risk Neutral Probability Densities by Fitting Implied Volatility Smiles: Some Methodological Points And An Application To The 3M Euribor Futurs Option Prices* Working Paper No. 198, Working Paper Series, ECB, 2002.

[3] Bartlett *Hedging under the SABR* Wilmott Magazine, 2006.

[4] Bauer and Rudebusch *The Signaling Channel for Federal Reserve Bond Purchases* Working Paper 2011-21, Federal Reserve Bank of San Francisco , 2013.

[5] Berger and al. *Monetary Policy in the Media* ECB Working Paper Series No 679 , 2006.

[6] Blake, A.P., Rule, G.R. and Rummel *Inflation targeting and term premia estimates for Latin America* O.J. Lat Am Econ Rev 24: 3 , 2015.

[7] Blei and al. *Latent Dirichlet Allocation* Journal of Machine Learning Research 3 993-1022 , 2003.

[8] Blei and Lafferty *Dynamic Topic Models* ICML '06 Proceedings of the 23rd international conference on Machine learning Pages 113-120 , 2006.

[9] Blei and Lafferty *Topic Models* , 2009.

[10] Blei and Lafferty *Probabilistic Topic Models* , Communications of the ACM, Vol. 55 No. 4, Pages 77-84, 2012.

[11] Blinder and al. *Central Bank Communication and Monetary Policy: a Survey of Theory and Evidence* , ECB Working Paper Series no 898, 2008.

[12] Bliss and Panigirtzoglou *Testing the stability of implied probability density functions*, 2000.

[13] Breeden and Litzenberger *Prices of State-Contingent Claims Implicit in Option Prices*, The Journal of Business, 1978.

[14] Carroll, Christopher D. *Macroeconomic expectations of households and professional forecasters*, Working Papers, The Johns Hopkins University, Department of Economics, 2003.

[15] Cox and Ross, *The Valuation of Options for Alternative Stochastic Processes*, Working Papers, Journal of Financial Economics 3 145-166, 1976

[16] DellaVigna, Stefano and Ethan Kaplan, *The Fox News Effect: Media Bias and Voting*, Quarterly Journal of Economics 122 , 2007.

[17] Ehrmann and Fratzscher, *Communication And Decision-Making by Central Bank Committees Different Strategies, Same Effectiveness?*, ECB Working Paper Series, No. 488, 2005.

[18] George W. Evans and Seppo Honkapohja, *Learning and Expectations in Macroeconomics*, Princeton University Press, 2001.

[19] George and Waldfogel, *The New York Times and the Market for Local Newspapers*, American Economic Review 96(1):435-447, 2006.

[20] Glick and Leduc, *Central Bank Announcements of Asset Purchases and the Impact on Global Financial and Commodity Markets*, Economic Research Department Federal Reserve Bank of San Francisco, 2011.

[21] Griffiths and Steyvers, *Finding scientific topics*, 2004.

[22] Grimmer and Stuart, *Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts*, 2012.

[23] Refet Gurkaynak and Brian P. Sack and Eric T. Swanson, *Do actions speak louder than words? the response of asset prices to monetary policy actions and statements*, Finance and Economics Discussion Series , 2004.

[24] De Haan and al. *The Transparency and Credibility of the European Central Bank*, Journal of Common Market Studies, 2004.

[25] James D. Hamilton *Time Series Analysis*, Princeton University Press, 1994.

[26] Stephen Hansen and Michael McMahon *Shocking language: Understanding the macroeconomic effects of central bank communication*, Journal of International Economics, 2016.

[27] Bernd Hayo and Edith *The German public and its trust in the ECB: The role of knowledge and information search*, Journal of International Money and Finance, 2014.

[28] Harrison Honga, Sven Rady, *Strategic trading and learning about liquidity*, Journal of Financial Markets, 2002.

[29] Lixin Huang, Hong Lui, *Rational Inattention and Portfolio Selection*, The Journal of Finance, 2007.

[30] Hubert, Paul *Central Bank Forecasts as an Instrument of Monetary Policy*, Documents de Travail de l'OFCE, 2012.

[31] David-Jan Jansen and Jakob de Haan *An Assessment of the Consistency of ECB Communication using Wordscores*, DNB Working Paper, 2010.

[32] Jung *Have monetary data releases helped markets to predict the interest rate decisions of the European Central Bank?*, ECB Working Paper, 2016.

[33] Kohn and Sack *Central Bank Talk: Does It Matter and Why?*, Board of Governors of the Federal Reserve System Washington, 2004.

[34] Michael Lamla and Sarah Lein *The role of media for consumers' inflation expectation formation*, Journal of Economic Behavior and Organization, 2014.

[35] Michael J. Lamla and Jan-Egbert Sturm *Interest Rates, ECB Communication and Inflation Expectations*, 2013.

[36] Laver and al. *Extracting Policy Positions from Political Texts Using Words as Data*, he American Political Science Review, 2003.

[37] Lucca and Trebbi *Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements*, 2011.

[38] Thomas Maag and Michael Lamla *The Role of Media for Inflation Forecast Disagreement of Households and Professionals*, KOF Working papers, 2009.

[39] MacCombs and Shaw *The Agenda-Setting Function of Mass Media*, The Public Opinion Quarterly, 1972.

[40] Malz *A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions*, Federal Reserve Bank of New York Staff Reports, 2014.

[41] N. Gregory Mankiw and Ricardo Reis *Sticky information versus sticky prices: a proposal to replace the New-Keynesian Phillips curve*, 2001.

[42] Stephen Morris and Hyun Song Shin *Social Value of Public Information*, The American Economic Review, 2002.

[43] Stephen Morris and Hyun Song Shin *Contagious Adverse Selection*, The American Economic Journal, 2012.

[44] Pereira das Neves *The Mass Media Transmission of Central Bank Communication under Uncertainty*, 2016.

[45] Rachel Reeves and Michael Sawicki *Do financial markets react to Bank of England communication?*, European Journal of Political Economy, 2007.

[46] Hubert and Maule*Policy and Macro Signals as Inputs to Inflation Expectation Formation*, Staff Working Paper No. 581, Bank of England, 2016.

[47] Reis, Ricardo, *Central Bank Design*, Journal of Economic Perspectives, 2013.

[48] Nachum Sicherman, and al. *Financial Attention*, The Review of Financial Studies, 2016.

[49] Sims, Christopher A. *Implications of rational inattention*, Journal of Monetary Economics, 2003.

[50] Sims, Christopher A. *Rational Inattention: Beyond the Linear-Quadratic Case*, American Economic Review, 2006.

[51] Tom Smith *Option-Implied Probability Distributions for Future Inflation*, Bank of England Quarterly Bulletin, 2012.

[52] Antonella Tutino *Rational Inattention? Guides Overloaded Brains, Helps Economists Understand Market Behavior*, Economic Letter vol. 6, 2001.

[53] Veronesi *Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model*, The Review of Financial Studies, 1999.

[54] Wainwright and Jordan, *A Variational Principle for Graphical Models*, Chapter 11 in New Directions in Statistical Signal Processing, MIT press, 2005.

[55] Walsh *Monetary Policy and Theory*, third edition, MIT Press, 2003.

[56] Werner F. M. De Bondt and Richard Thaler *Does the Stock Market Overreact?*, The Journal of Finance, 1984.

[57] Michael Woodford *Central Bank Communication and Policy Effectiveness*, NBER Working Paper, 2005.