

Paloviita, Maritta; Haavio, Markus; Jalasjoki, Pirkka; Kilponen, Juha; Vänni, Ilona

**Working Paper**

## Reading between the lines: Using text analysis to estimate the loss function of the ECB

Bank of Finland Research Discussion Papers, No. 12/2020

**Provided in Cooperation with:**

Bank of Finland, Helsinki

*Suggested Citation:* Paloviita, Maritta; Haavio, Markus; Jalasjoki, Pirkka; Kilponen, Juha; Vänni, Ilona (2020) : Reading between the lines: Using text analysis to estimate the loss function of the ECB, Bank of Finland Research Discussion Papers, No. 12/2020, ISBN 978-952-323-337-9, Bank of Finland, Helsinki, <https://nbn-resolving.de/urn:nbn:fi:bof-202007032225>

This Version is available at:

<https://hdl.handle.net/10419/240335>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Bank of Finland Research Discussion Papers  
12 • 2020

Maritta Paloviita – Markus Haavio – Pirkka Jalasjoki –  
Juha Kilponen – Ilona Vänni

Reading between the lines – Using text analysis  
to estimate the loss function of the ECB



Bank of Finland  
Research

Bank of Finland Research Discussion Papers  
Editor-in-Chief Esa Jokivuolle

Bank of Finland Research Discussion Paper 12/2020  
6 July 2020

Maritta Paloviita – Markus Haavio – Pirkka Jalasjoki – Juha Kilponen – Ilona Vänni:  
Reading between the lines – Using text analysis to estimate the loss function of the ECB

ISBN 978-952-323-337-9, online  
ISSN 1456-6184, online

Bank of Finland  
Research Unit

PO Box 160  
FIN-00101 Helsinki

Phone: +358 9 1831

Email: [research@bof.fi](mailto:research@bof.fi)

Website: [www.suomenpankki.fi/en/research/research-unit/](http://www.suomenpankki.fi/en/research/research-unit/)

The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of Finland.

# Reading between the lines – Using text analysis to estimate the loss function of the ECB

Maritta Paloviita\* – Markus Haavio – Pirkka Jalasjoki  
– Juha Kilponen – Ilona Vänni

July 3, 2020

## Abstract

We measure the tone (sentiment) of the ECB’s Governing Council regarding economic outlook at the time of each monetary policy meeting and use this information together with the Eurosystem/ECB staff macroeconomic projections to directly estimate the Governing Council’s loss function. Our results support earlier, more indirect findings, based on reaction function estimations, that the ECB has been either more averse to inflation above 2% ceiling or that the *de facto* inflation aim has been considerably below 2%. Our results suggest further that an inflation aim of 2% combined with asymmetry is a plausible specification of the ECB’s preferences.

**JEL Codes:** E31, E52, E58

**Keywords:** central bank communication, ECB, monetary policy, textual analysis, inflation target, real time projections, loss function

---

\*Corresponding author: email [maritta.paloviita@bof.fi](mailto:maritta.paloviita@bof.fi). All authors are from the Bank of Finland. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of Finland or the Eurosystem.

# 1 Introduction

The quantitative definition of price stability is the main building block of the European Central Bank's (ECB) monetary policy strategy, affecting monetary policy discussion and decision making in the ECB's Governing Council (GC). It also affects the way the central bank communicates to the public. A thorough analysis of the quantitative definition of price stability is a crucial part of the strategy review work in the Eurosystem in 2020 - 2021.

The ECB's monetary policy is guided by a two-step quantification of price stability by the GC. In 1998, the GC defined price stability as a '*year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2%*'. In 2003, the GC clarified that '*in the pursuit of price stability it aims to maintain inflation rates below, but close to, 2% over the medium term*'.

The two-step definition of price stability may have obscured what the inflation aim of the ECB is, and its theoretical rationale is unclear. The *de facto* inflation aim of the ECB has been interpreted in various ways both inside and outside the ECB. In particular, the expression '*below, but close to 2%*' has some feel of asymmetry even though the ECB's communication stresses symmetry in its monetary policy making. Ambiguity of the price stability definition may hamper internal work of the GC and can lead to policy mistakes.

Rostagno et al. (2019) discuss the narrative of the ECB's monetary policy during the past 20 years. They argue that during the early years of the monetary union, the price stability range may have been too wide to steer monetary policy appropriately. Policy decisions were rationalized by medium term upside or downside risks to price stability rather than by inflation outcomes relative to the GC's inflation aim. This may have hampered monetary policy communication, increasing risks of de-anchoring inflation expectations and leading to unnecessary policy uncertainty. Ambiguity of the inflation aim may have jeopardized the GC's ability to direct monetary policy appropriately also in the post-crisis years, when inflation pressures were low due to shortfalls in demand, and monetary policy was constrained by the zero lower bound. At that time also policy trade-offs became more complex and uncertainty over future course of the economy increased.

The aim of this study is twofold, as we assess what the level of the *de facto* inflation aim of the ECB is, and as we examine whether the ECB has been more averse to high inflation than to low inflation. We use text mining techniques (natural language processing) in order to “read between the lines” and infer the ECB’s preferences, i.e. its perceived discontent to economic developments (loss) directly by focusing on the introductory statements of the ECB’s press conferences. Applying the dictionary-based “bag of words” approach we use Loughran and McDonald’s (2011) finance-specific dictionary in order to measure a latent component, i.e. the general tone (sentiment) in the ECB’s qualitative communication. We measure net negativity, i.e. the difference between the fraction of negative words and phrases and fraction of positive words and phrases in each introductory statement. Then, following Shapiro and Wilson (2019) – who study the Federal Open Market Committee – we use the tone together with the HICP inflation and output gap projections of the Eurosystem/ECB staff to directly estimate the ECB GC’s loss function. We estimate both piecewise linear and more general Linex (Linear-exponential) types of loss functions. A Linex function encompasses a quadratic loss function, which is the most usual formulation of central bank preferences in a theoretically oriented optimal monetary policy literature. In the end, we repeat the estimations using an alternative proxy for the tone, which is based only on inflation-related texts in the introductory statements. The alternative proxy is constructed using the Latent Dirichlet Allocation (LDA) method.

Our results are in line with earlier findings by Hartmann and Smets (2018), Paloviita et al. (forthcoming) and Rostagno et al. (2019) based on reaction function estimations. They suggest that either the targeted inflation rate has been relatively low (1.6% – 1.8%), or that the ECB has targeted inflation rates close to the 2% ceiling, but the policy responses to inflation rates above the target have been stronger than to inflation rates below the target (asymmetric policy). Our analyses indicate that if we assume a priori that the ECB has conducted symmetric monetary policy, estimates of the *de facto* inflation target of the ECB are relatively low (1.7%). If, however, we fix the *de facto* inflation target to the upper bound of the price stability definition (2%), the loss function estimation reveals asymmetric preferences to inflation. These results are robust to inclusion or exclusion of secondary targets (such as output gap or squared output gap) to the loss function and also to whether we estimate a piecewise linear or a more general Linex loss function. Similarly to the earlier research, it is not easy to separate between a low *de facto* inflation target and asymmetry. However,

compared to earlier research we find somewhat stronger evidence in favor of asymmetric preferences.

Our study contributes to the existing literature in several ways. First, it relates to studies which use text mining techniques to study qualitative communication of monetary policy. This relatively new technique in economics refers to a large group of computational tools and statistical methods to quantify text, i.e. mapping text to some meaningful quantities. Contrary to human reading, a huge amount of written material can be analysed without any prior assumptions using automated text mining methods. Bholat et al. (2015) present a comprehensive summary of these techniques applied to central bank documents.<sup>1</sup>

In a number of recent studies textual analysis methods have been applied in order to study monetary policy communication. For example, Baranowski et al. (2020) examine the ECB's introductory statements and construct a measure for the ECB's tone by applying a bag-of-words approach and several dictionaries. When studying the stance of the ECB's monetary policy and the state of the euro area economy according to the ECB's introductory statements, Picault and Renault (2017) use an approach which is based on term-weighting and contiguous sequences of words (n-grams). Berger et al. (2011) construct an index which summarizes the favorableness with which the ECB's monetary policy decisions are discussed in the media. The focus of Ehrmann and Talmi (2020) is mainly on the of the Bank of Canada's press releases. Using Hansen and McMahon's (2016) dictionary, they examine how semantic changes in subsequent central bank press releases affect market volatility. Armelius et al. (2020) apply Loughran and McDonald's (2011) dictionary to conduct a sentiment analysis of speeches from 23 central banks. The Bank of Japan's Monthly Reports are investigated by Kawamura et al. (2019) and the Bank of England's Inflation Reports are studied by Jones et al. (forthcoming). Kawamura et al. (2019) analysis is based on the Japanese Sentiment Polarity Dictionary and a Latent Dirichlet Allocation (LDA) model, whereas the focus in Jones et al. (forthcoming) is on score analysis. Bennani et al. (2020) construct a measure for communication of the ECB's GC members between policy meetings. Their measure reveals whether inter-meeting verbal communication refers to easing, tightening or maintaining the monetary policy stance. Korhonen and Newby (2019) explore tweets by 46 European central banks and financial supervisors.

---

<sup>1</sup>For an introduction to text mining techniques, see Manning et al. (2008).

Perhaps the closest study to ours is the recent paper by Shapiro and Wilson (2019), who estimate objective functions of the U.S. Federal Open Market Committee (FOMC). They link in their estimations the proxy for the FOMC's loss (net negativity) to the Federal Reserve's staff Greenbook forecasts of the PCE inflation and real economic variables as well as contemporaneous stock market variables. Their main analysis is based on transcripts of the FOMC meetings, but they also study public speeches of individual FOMC members and minutes of the FOMC meetings. Shapiro and Wilson (2019) find that the FOMC's implicit inflation target has been about 1.5% in 2000 to 2013 and the FOMC's loss depends clearly on output growth and stock market performance, but it is not closely related to the FOMC's perception of the current economic slack.

Second, our paper contributes to the literature on real time information and central bank forecasting which are essential in monetary policy making. Many authors have recently examined the impact of central bank forecasts on private sector forecasts and expectations formation (see e.g. Fujiwara (2005), Hubert (2014, 2015, 2017) and Lyziak and Paloviita (2017, 2018)). Also the accuracy of published forecasts has been studied quite intensively in many studies (Potter (2011), Stockton (2012), Fawcett et al. (2015), Iversen et al. (2016), Kontogeorgos and Lambrias (2019)). Our analysis is based on a unique data set, which includes real time quarterly macroeconomic projections of the Eurosystem/ECB staff. The same information is presented to the GC in monetary policy meetings. Using Shapiro and Wilson's (2019) approach, we are able to link the proxy of the ECB's loss to the ECB's key policy objectives by estimating the loss function of the ECB.

Our approach of estimating the loss function directly complements the earlier, more indirect analyses based on estimated reaction functions in a novel way. Contrary to a reaction function based approach presented by Hartmann and Smets (2018), Paloviita et al. (forthcoming) and Rostagno et al. (2019), we do not have to make specific assumptions about the functional form of the reaction function, the structure and stability of the economic relationships or the transmission channels of monetary policy, in order to infer the parameters of the loss function. A text mining approach has also additional advantages due to the fact that the dependent variable is the tone (sentiment) proxy instead of the policy instrument itself or policy-sensitive short-term market interest rates. When approaching the zero or negative effective lower bound and when introducing non-standard monetary policy measures, the usual linear policy reaction function breaks down



(or alternatively, rather uncertain proxies, i.e. shadow rates must be used). Using the tone proxy as the dependent variable, the estimation period can be extended to periods of unconventional measures without a fear of non-linearity or discontinuity in the reaction function. Furthermore, in reaction function estimations using a policy instrument itself, the reaction function is identified only from incremental interest rate changes, measured in 0.1 or 0.25 percentage points, and not from a continuous mapping between economic outcomes and preferences.

The paper is organised as follows. Central bank communication, methodology and data are discussed in sections 2 and 3. Estimation of the loss function is presented and empirical analysis is reported in sections 4 and 5 and concluding remarks are provided in section 6.

## 2 Central bank communication and general tone

Central banks possess considerable amount of information on economic conditions, which goes beyond that available to the private sector.<sup>2</sup> Central banks have a clear incentive to disclose such information since it can be useful for the private sector to make more informed economic and financial decisions and for the central bank to achieve its objectives more efficiently by steering the private sector expectations on interest rates and inflation to the direction it prefers.<sup>3</sup> A central bank's communication can be both quantitative (forecast numbers) and qualitative (formal statements/reports, informal speeches/interviews and social media such as Twitter). A good example of purely qualitative but powerful communication is the ECB President Draghi's speech in London in July 2012. In that speech, his statement that *'within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough'*<sup>4</sup> practically ended the discussion on the break-up of the euro area, and led to a massive re-pricing of euro area sovereign bonds. Another good example of both qualitative and quantitative communication is forward guidance, where the central bank announces its commitment to a certain path of interest rates and links this decision to its public statements

---

<sup>2</sup>See e.g. Romer and Romer (2000).

<sup>3</sup>See e.g. Blinder et al. (2008) and Haldane and McMahon (2018) for an extensive discussion on central bank communication.

<sup>4</sup><https://www.ecb.europa.eu/press/key/date/2012/html/sp120726.en.html>

about current and projected economic conditions.<sup>5</sup>

The introductory statements of the press conferences are the most important form of formal communication of the ECB. Contrary to, for example, speeches and interviews, the content of the introductory statements is always highly relevant from the point of view of monetary policy stance, as they include carefully drafted announcements and explanations of the GC's decisions based on real time economic and monetary analysis of the ECB/ESCB staff. The introductory statements always include a summary of monetary policy decisions and economic analysis including short to medium term outlook. They also include monetary analysis, a cross-check paragraph combining conclusions from the economic and monetary analysis and sections focusing on fiscal policy issues and structural reforms. After every ECB press conference, key phrases and semantic changes in the introductory statement are intensively analysed by media and financial market participants in order to gain information about the central bank's policy objectives, future policy actions and its assessment of current and future economic conditions. For example, in December 2019, Martin Arnold assessed in the Financial Times that '*Lagarde follows Draghi's policy lead but sounds [an] upbeat tone*'.<sup>6</sup> The introductory statements include also the Eurosystem/ECB staff macroeconomic projections (quantitative communication) four times a year.

There are various ways to construct a proxy for the sentiment of qualitative communication and to measure an overall tone within the text. The most common techniques are based on traditional lexicon-based approaches, which use special dictionaries of words annotated with their semantic orientation. Semantic orientation can be understood as a measure of subjectivity and opinion, capturing positivity, negativity (or neutrality) and strength (degree to which the text is positive or negative) on some topic, person, or idea (Osgood, Suci, and Tannenbaum, 1957). Advantage of lexicon-based methods is simplicity, as they do not require labelled training data. Machine learning techniques, especially Support Vector Machines (Sharma and Dey, 2012) and artificial neural networks (Moreas et al., 2013), are superior to lexicon-based techniques in terms of

---

<sup>5</sup>Fujiwara and Waki (2019) study welfare benefits of forward guidance in a setting where the central bank possesses exclusive information on economic conditions relative to the private sector.

<sup>6</sup>In July 2019, Carsten Brzeski noted that '*the tone of the minutes reflects the ECB's concerns about the growth and inflation outlook*' (<https://think.ing.com/snaps/ecb-draghi-tltro-rates-qe-ing-carsten-brzeski-talk-june-minutes-meeting/>). In January 2020, the title of David Goodman's article was '*Trio of G-7 Central Banks May Set Tone for 2020*' (<https://www.bloomberg.com/news/articles/2020-01-19/trio-of-g-7-central-banks-may-set-tone-for-2020-eco-week-ahead>).

classification accuracy especially when handling domain-specific language (e.g. central bank specific) rather than general language (Kumar and Jaiswal, 2016). On the other hand, automatic polarity classification requires only a minor effort in terms of manual annotation.

A tone index has been defined in the literature in different ways. For example, Hansen and McMahon (2016) measure the tone in the FOMC’s communication as a net positivity using the fractions of positive and negative words. Schmeling and Wagner (2019) also analyze communication of the FOMC, but they consider only the fraction of negative words and ignore negations. Loughran and McDonald (2016) argue that with this approach the consideration of negations is not that critical.

### 3 Methodology and data

We extract the sentiment or tone from the ECB’s introductory statements by means of text analysis. We have gathered press conference texts from the ECB’s website<sup>7</sup> using automated web scraping. Our sample covers the period from July 1999 to December 2019. At the beginning of the sample, until December 2014, the GC had a monetary policy meeting every month, but after that the GC has made monetary policy decisions eight times a year.

First, we have pre-processed the introductory statement data using the most common techniques in textual analysis. To reduce noise in the data, we have merged some common word sequences into single terms (e.g. ‘governing council, ‘european central bank’) and deleted paragraphs of less than 10 words. These short paragraphs mainly include only general expressions and thus do not concern economic analysis. We have also deleted first paragraphs of each introductory statement, as they represent greetings (‘Ladies and gentlemen...’). The expression ‘let me now explain our assessment in greater detail’ has been explicitly excluded, since this sentence appears often and would be misclassified as positive with automatic classification. We have also converted upper cases into lower cases, removed punctuation, numbers, extra whitespaces and common English stopwords excluding negations, and finally stemmed the words.

---

<sup>7</sup><https://www.ecb.europa.eu/press/pressconf/html/index.en.html>

In the next step, data are split into tokens (single words). Paragraphs are at first treated as separate documents (data points), but then documents are aggregated into a quarterly level. Even though it is relatively complicated to take into account the grammatical relations of terms when using a lexicon-based classification, we still consider negations as their impact on the overall tone is potentially significant. Since it is reasonable to assume that central banks do not announce good news via negative expressions combined with negations, e.g. ‘confidence has not deteriorated’, but instead bad news may be softened by expressing negative messages using positive words (‘confidence has not improved’), we take into account only negations followed by positive words as suggested by Loughran and McDonald (2011). To handle negations, we treat each sentence separately. If a sentence consists of a negation and a positive word within three words after the negation, the sentiment of the word is reversed into negative, as suggested by Loughran and McDonald. As shown in Figure 1, the average length of introductory statements is 647 words, and the longest introductory statements are observed in the middle of the sample (2006Q4 – 2011Q4).

Finally, to measure the general tone in the ECB’s introductory statements, we use polarity classification to classify words into three classes (negative, positive and neutral) based on their sentiment. For classification, we use Loughran and McDonald’s (2011) finance-dictionary, which we modify further by adding some central bank specific terms and expressions in order to improve accuracy.<sup>8</sup> We count negative and positive words found in the modified Loughran and McDonald dictionary and define the tone as a difference of the number of negative and positive words, normalized with the total number of words. We use the tone as a proxy of loss, as we want to measure negativity when estimating loss functions. Consequently, the measured tone index gets larger values as negative sentiment increases, corresponding to a net negativity percentage of the total corpus. More specifically, we follow Shapiro and Wilson (2019) and determine the tone index as

$$N_t = \frac{\#Neg - \#Pos}{\#Tot}. \quad (1)$$

See Appendix 1 for an example of scored text. Figure 2 shows the evolution of the tone, i.e. net negativity

---

<sup>8</sup>Added positive words: ‘ample’, ‘favourable’, ‘restructuring’, ‘robust’, ‘buoyant’, ‘dynamic’. Added negative words: ‘dis-equilibria’, ‘deficit’, ‘tensions’. Added expressions: ‘unemployment’ + ‘decline’/‘fall’/‘decrease’, ‘risk to price stability’ + ‘low’/‘high’/‘increase’/‘decrease’.

in the introductory statements of the ECB. Broadly speaking, the net negativity in the ECB’s introductory statements gradually decreased in the pre-financial crisis years, but the financial crisis contributed to a huge increase in the net negativity, as may be expected. In the post-financial crisis years, a gradual decrease in the net negativity is observed until 2017Q4, but recently the tone has been more negative again. On average, the net negativity has been slightly negative in the whole sample. Overall, movements in the tone are clearly associated to economic developments in the euro area. In estimations, we use the tone as a dependent variable and hence it is a proxy for the short term loss of the ECB.

## 4 Estimation of loss functions

In this section, we attempt to quantify how the GC’s definition of price stability is reflected in its policy preferences. We estimate the parameters of a loss function where we link the tone in the introductory statements to deviation of inflation from some *de facto* inflation aim. Our choice to focus on inflation as a key determinant of the loss reflects the fact that price stability is the primary objective of the ECB’s monetary policy. However, we also control for a measure of output gap as an additional determinant of the loss.<sup>9</sup>

In the first step, we estimate piecewise linear specifications. While piecewise linear functions are statistically less demanding to capture possible asymmetries, the caveat is that they do not allow for convexity in the preferences; the loss increases linearly with respect to a distance from, say, an inflation target. Therefore, in the second step we estimate linear-exponential (i.e. Linex) loss functions which are more flexible, as they nest both quadratic and asymmetric (convex) preferences. For example, Surico (2003) estimates Linex loss functions for the ECB and Surico (2007) for the Federal Reserve. The same approach is applied for the UK and Canada by Caglayan et al. (2016) and for Australia and New Zealand by Karagedilkli and Lees (2004). Asymmetric quadratic losses for the Federal Reserve are examined by Capistran (2008).

---

<sup>9</sup>Shapiro and Wilson (2019) add also different measures of real economic activity, such as output gap, output growth, unemployment (gap), and financial variables into the right hand side of the regression.

## 4.1 Piecewise linear loss function

The simplest way to estimate the parameters of a loss function is to assume a V-shaped, i.e. a piecewise linear functional form. A V-shaped loss function is in line with the idea that the central bank has a preference to keep inflation close to its inflation aim and it does so by minimizing the deviation, i.e. distance of actual inflation from the targeted level of inflation subject to some constraints and trade-offs. A V-shaped loss function is not restricted to be symmetric such as a quadratic loss function, which is a typical (but not necessarily the most realistic) assumption in the traditional models of central bank preferences and optimal monetary policy (see e.g. Woodford (2003), Walsh (2003), and Gali (2015)).

We start with a general short run loss function expressed as follows:

$$L_t = |\tilde{\pi}_t|. \tag{2}$$

In equation (2), the inflation gap, i.e. the term  $\tilde{\pi}_t = \pi_t - \pi^*$  is defined as the difference between inflation  $\pi_t$  and the inflation aim  $\pi^*$ . We proxy the central bank loss by the tone, i.e. we assume that there is a linear correspondence between the tone and the central bank's loss function such that  $N = \alpha + \delta L$ .<sup>10</sup> As explained above, the tone  $N$  measures the net negativity in the introductory statements of the ECB's press conferences. We add an i.i.d. error term in order to take into account measurement error in the loss and to relate the tone into inflation outcomes (and forecasts) in the following way:

$$N_t = \alpha + \delta |\tilde{\pi}_t| + \varepsilon_t. \tag{3}$$

Next, we split the linear loss function into two separate segments using a dummy variable. The dummy  $D$  is equal to zero if the real time estimate of inflation is below the inflation aim of the central bank (i.e. the inflation gap is negative). Correspondingly, the dummy  $D$  is equal to one if the real time estimate of inflation rate is above the inflation aim of the central bank (i.e. the inflation gap is positive). This gives rise to a more general loss function

---

<sup>10</sup>Parameter  $\alpha$  is simply a regression constant.

$$N_t = \alpha + \delta_B \tilde{\pi}_t (1 - D) + \delta_A \tilde{\pi}_t D + \varepsilon_t. \quad (4)$$

First, we use the constraint  $\delta_B + \delta_A = 0$  to estimate a symmetric loss function. Then, we repeat the estimation without this constraint in order to allow asymmetry, i.e. so that the parameters  $\delta_B$  and  $\delta_A$  can differ from each other. Later on, we add also terms of output gap  $\Delta \tilde{y}_t$  and squared output gap  $\Delta \tilde{y}_t^2$  on the right hand side such that the most general form of the loss function reads as

$$N_t = \alpha + \delta_B \tilde{\pi}_t (1 - D) + \delta_A \tilde{\pi}_t D + \beta_1 \Delta \tilde{y}_t + \beta_2 \Delta \tilde{y}_t^2 + \varepsilon_t. \quad (5)$$

To measure inflation at the time of decision making, we use the Eurosystem/ECB inflation projections presented to the GC in the monetary policy meetings. Macroeconomic projections of the Eurosystem/ECB staff, which include all real time economic information available at the time, contain information on both the short run and long run economic outlook. As emphasized by Shapiro and Wilson (2019), Romer and Romer (2004) and Coibion and Gorodnichenko (2011), short run economic forecasts can be treated as exogenous to policy since monetary policy impacts on the economy only with some delay. Furthermore, the economic outlook is prepared and finalized by the Eurosystem/ECB staff before the GC's policy meetings.

In what follows, we focus on short run inflation forecasts when estimating the loss function. We proxy the inflation perception of the ECB ( $\pi_t$ ) by the average of the nowcast inflation rate and one-quarter-ahead projected inflation rate, following Shapiro and Wilson (2019). We measure output gap  $\Delta \tilde{y}_t$  by the difference between the average of the nowcast and one-quarter-ahead output growth projections and potential output growth, which is proxied by the medium term (eight-quarters-ahead) GDP growth projection. The underlying assumption is that the medium run growth projection for the euro area corresponds to the assessed real time euro area growth potential.<sup>11</sup> The introductory statements have been released once a month or eight times a year. Since our estimations are based on the quarterly inflation projections of the Eurosystem/ECB staff, we calculate the net negativity ( $N_t$ ), i.e. the tone, by combining all the text in the introductory statements

---

<sup>11</sup>Another option would have been to use potential output estimates. However, the real time estimates for euro area potential output are only available from 2009Q2 at a quarterly frequency and from 2006 at an annual frequency in the ECB projection data. It is also worth noting that in the ECB's New Area-Wide Model (NAWM), the reaction function has been specified in terms of deviations of output growth from its long-run empirical mean (Christoffel et al. 2008).

held during each quarter.

We estimate  $\pi^*$  by doing a grid search for a linear Ordinary Least Squares (OLS) regression model. To obtain  $\pi^*$ , we first search for the best estimate by stepping through all possible values of  $\pi^*$  between 0 and 3 with a step size of 0.01. Then we choose those specifications that minimize the sum of squared residuals (SSR) in equations 3 to 5. See Appendix 2 for the corresponding SSR plots. From the chosen specifications, we report estimation results in Table 1. In the case of the symmetric specification, we obtain the confidence interval for  $\pi^*$  using a bootstrap procedure. Using the F-test and likelihood ratio tests, we compare the empirical relevance of the unrestricted (asymmetric) and restricted (symmetric) specifications of the V-shaped loss function.

## 4.2 Results from piecewise linear loss function

Estimation results are summarized in Table 1. In the first step, we assume that the ECB's monetary policy responses to negative and positive inflation gaps are equal, and estimate a symmetric piecewise linear loss function, i.e. Equation 3. For this specification, the estimated slope parameter  $\hat{\delta}$  is somewhat above one and the point estimate for the implicit *de facto* inflation aim is 1.66% (see column (1)). The 95% confidence interval of the estimated inflation aim is relatively tight, ranging from 1.30 to 1.95. The positive slope coefficient is statistically highly significant and it suggests that as inflation (the average of HICP inflation nowcast and one-quarter-ahead projection) has been further away from 1.66%, the net negativity in the tone (loss) has increased considerably. Adding the squared and linear output gap terms does not change the results, and in fact, the point estimate for the *de facto* inflation target remains the same (see column (2) in Table 1).

In the second step, we relax the restriction on a symmetric loss function and estimate Equation 4, i.e. including separate parameters for negative and positive inflation gaps, but we fix the inflation aim to 1.66% (see columns (3) and (4)). According to column (3), the estimated coefficients  $\hat{\delta}_B$  and  $\hat{\delta}_A$  (-1.039 and 1.253, respectively) are somewhat different in absolute terms, but still relatively close to the corresponding point estimate in the symmetric loss function (1.104). Adding the output gap terms into the loss function yields



more distinct slope coefficients, pointing to more asymmetric preferences as regards to inflation outcomes (column (4)). The F-test now rejects symmetry at 8% significance level.

In the third step, in columns (6) and (8), we consider how the loss function is changed if the *de facto* inflation aim of the ECB is assumed to be at the upper bound of the price stability definition. We set exogenously the inflation target equal to 2.0% and re-estimate an unrestricted (i.e. possibly asymmetric) specification of the loss function (i.e. equations 3 and 4). In this case, the difference between the estimated slope coefficients is substantial (-0.634 and 1.529, respectively, in column (6) and -0.309 and 1.600, respectively, in column (8)). These parameter estimates suggest that when inflation has increased further above 2%, the net negativity in the tone has increased  $2\frac{1}{2}$  times or 5 times faster than when inflation has been falling further below 2%, depending on whether the estimation excludes or includes the output gap terms. This is the most direct evidence on the ECB being more averse to inflation rates above 2% than below 2% in our dataset. Finally, we have estimated an asymmetric specification without fixing in advance the inflation aim of the ECB. Columns (5) and (7) of Table 1 indicate that without (with) the output gap terms, the point estimate of the implicit *de facto* target of the ECB stands at 1.7% (2.3%). When interpreting the results from the estimation with the output gap terms (column (7)) it is worth noting that the sum of squared residuals (SSR) criterion we use for selecting the estimate of the *de facto* target is more or less indifferent between any values over the range 2.0% to 2.3% (see Figure A2.4 in the Appendix). The estimated loss functions without the output gap terms are presented graphically in Figures 3 to 5.

We use the F-test and likelihood ratio test in order to assess how the targeted inflation rate and possible asymmetry of the loss function are related. We compare symmetric and asymmetric specifications by taking the target inflation rate (1.66%) obtained from the specifications in columns (1) and (2), and fixing it to be used in the specifications of columns (3) and (4), which allow for asymmetry. The p-value of the F-test indicates that the specification in column (3), which combines a relatively low targeted inflation rate and asymmetry, does not lead to a significantly improved fit over the corresponding symmetric (restricted) model in column (1). Once the output gap terms are included in the regression (when comparing columns (2) and (4)), the F-test rejects the assumption of symmetry at an 8% significance level. The null hypothesis behind the F-test is close to being rejected also in the case when the inflation aim is fixed to 2.0% in column (6) and

is strongly rejected in column (8), in which the specification includes the output gap terms. At the same time, the likelihood ratio test does not reject the restriction of setting the *de facto* inflation rate exogenously to 2.0%.<sup>12</sup> This suggests that, in terms of statistical criteria, a 2.0% inflation aim combined with asymmetry is, to an extent, a plausible specification of the ECB's preferences.

### 4.3 Linex loss function

We estimate linear-exponential (Linex) loss functions by assuming that the ECB's preferences for inflation can be expressed in the following form:

$$L_t = \frac{\exp[\theta(\pi_t - \pi^*)] - \theta(\pi_t - \pi^*) - 1}{\theta^2}. \quad (6)$$

Here the parameter  $\theta$  captures asymmetry and convexity of the preferences (with respect to deviations of inflation from the target). Inflation outcomes above the target are increasingly more costly when  $\theta$  is large, while as  $\theta$  approaches zero, preferences take a quadratic form. Once again we proxy the central bank loss by the tone, and we assume that there is a linear correspondence between the tone and the central bank's loss function such that  $N = \alpha + \gamma L$ . Hence we relate the tone to inflation deviation in the following way:

$$N_t = \alpha + \gamma \frac{\exp[\theta(\pi_t - \pi^*)] - \theta(\pi_t - \pi^*) - 1}{\theta^2} + \varepsilon_t. \quad (7)$$

Nobay and Peel (2003) show theoretically, in a simple static optimal monetary policy model, that a central bank's asymmetric preferences (such as those presented in equation 6) can lead to inflationary or deflationary bias relative to the inflation target.<sup>13</sup> Deflationary bias arise when the central bank is more averse to high

<sup>12</sup>The likelihood ratio test compares the restricted model (with a pre-fixed inflation aim) to the unrestricted model with the same set of explanatory variables. Hence we compare the restricted specification in column (6) ((8)) to the unrestricted specification in column (5) ((7)).

<sup>13</sup>See also Orphanides and Wilcox (2003), Ruge-Murcia (2007) and Cassou et al. (2012) for monetary policy and inflation biases in models with non-quadratic preferences.

inflation. More specifically, they show that when the central bank optimizes subject to the above preferences and the Lucas supply curve<sup>14</sup>  $y = \beta(\pi_t - E\pi) + \epsilon_t$ , expected (average) inflation is

$$E\pi = \pi^* - 0.5 * \theta \sigma_\pi^2, \quad (8)$$

where  $\sigma_\pi^2$  is the variance of conditionally normally distributed inflation. The above equation says that inflation expectations fall below the target in particular when the central bank is strongly averse to high inflation (i.e. high  $\theta$ ) and when inflation is volatile. When preferences are quadratic (and symmetric), the expected inflation equals the target.

#### 4.4 Results from Linex loss function

In Table 2, we report the estimates of the parameter  $\theta$  and *de facto* inflation target  $\pi^*$ , together with the estimates of the parameters  $\alpha$  and  $\gamma$ , based on a linear-exponential loss function, with and without controlling for output gap and squared output gap as in Table 1. The estimated loss functions are presented graphically in Figures 6 and 7. The results are very much consistent with the findings using the simpler approach of a piecewise linear loss function. When we estimate all the parameters (including  $\pi^*$ ) jointly, the *de facto* inflation target is in the neighborhood of 1.7 and  $\theta$  is positive, but not statistically different from zero (see column (4), Table 2). The same is true for  $\theta$  when we fix the *de facto* inflation target to 1.66 or 1.70 (see columns (1) and (2) in Table 2). On the contrary, when we fix the *de facto* inflation target to 2.0, the parameter  $\theta$  is significantly different from zero, pointing to strong asymmetry in central bank preferences (column (3)). When we include the output gap and estimate the specification

$$N_t = \alpha + \gamma \frac{\exp[\theta(\pi_t - \pi^*)] - \theta(\pi_t - \pi^*) - 1}{\theta^2} + \beta_1 \Delta \tilde{y}_t + \beta_2 \Delta \tilde{y}_t^2 + \epsilon_t, \quad (9)$$

---

<sup>14</sup> $y$  denotes output,  $\beta$  is the slope of the Lucas supply function, and  $\epsilon$  an exogenous supply shock.

where  $\beta_1$  and  $\beta_2$  are parameters, we obtain similar results. When we estimate all the parameters jointly, the point estimate of the *de facto* inflation target is 1.63;  $\theta$  is positive, but not statistically different from zero (see column (8), Table 2). When we fix the *de facto* inflation target to 2.0, the parameter  $\theta$  is significantly different from zero (column (7) in Table 2), suggesting asymmetric preferences with respect to inflation.

Finally, applying the likelihood ratio test, we cannot reject the hypothesis that the ECB’s preferences are characterized by a *de facto* target of 2.0%, combined with strong aversion to inflation above the target (see columns (3) and (7) in Table 2), pointing to asymmetry.<sup>15</sup> This asymmetry can visually be seen particularly in Figure 7, which plots the estimated Linex loss function: the net negativity increases strongly for inflation rates above 2%, and clearly less so when inflation rates are below 2%. Using the simple formula for expected inflation in Equation 8, the point estimate for the parameter  $\theta$  from column (3) in Table 2, and the realized variance of inflation during 1999M1-2020M1, expected inflation would be equal to  $E\pi = 2 - 0.5 * 1.15 * 0.83 = 1.52$ , remarkably close to the actual average inflation over the same period (1.70).

## 5 Alternative tone

The analysis above is based on direct estimation of the loss function. In the estimations, we proxy the ECB’s preferences (i.e. its perceived discontent to economic developments) by the general tone, which is constructed from the ECB’s introductory statements using Loughran and McDonald’s (2011) finance-specific dictionary. This “general tone” includes carefully drafted information of the GC’s decisions based on real time information and it reveals the GC’s real time assessment of an appropriate monetary policy stance, given the ECB/Eurosystem staff’s view on the economic developments.

We repeat our estimations using an alternative proxy for the ECB’s tone. Instead of measuring the general tone based on all text in the introductory statements, we focus on a subsample of the total corpus associated only with inflation, the primary objective of monetary policy of the ECB.

---

<sup>15</sup>The likelihood ratio test compares the restricted model (with prefixed inflation aim) to the unrestricted model with the same set of explanatory variables. Hence we compare the restricted specification in column (3) ((7)) to the unrestricted specification in column (4) ((8)).

In this exercise, we combine probabilistic topic modelling and dictionary-based sentiment analysis. First, we extract paragraphs of the introductory statements concerning inflation by employing Latent Dirichlet Allocation (LDA) introduced by Blei et al. (2003). Then, we calculate the tone (net negativity) based on these inflation relevant texts. Finally, we estimate loss functions, which link the alternative tone to the Eurosystem/ECB staff projections of the HICP inflation.

LDA is a hierarchical Bayesian model for dimensionality reduction that summarizes documents as a mixture of topics and topics as a mixture of words. The basic idea of LDA is to define a probability for each word in a corpus of being generated from a specific topic and to define each document as a distribution over a collection of topics. As a result, each word  $w$  in a document  $d$  is allocated into topic  $k$  based on topics in a document and how frequently  $w$  appears in a particular topic  $k$ . LDA model assumes exchangeability and thus is a bag-of-words method as a lexicon-based sentiment analysis. One benefit of using such unsupervised clustering for topic classification is the reduction in manual work as documents do not need to be read by humans. Even though, as a Bayesian model, LDA requires some subjective decisions when defining priors, subjectivity is still reduced compared to manual classification since it is not predefined which words describe each topic.

LDA has recently become a common technique to cluster textual data among authors examining central bank communication. Using LDA, Hartmann and Smets (2018) find 50 separate topics in the ECB's speeches over time, while Tobback et al. (2017) utilize LDA to extract topics in the news articles concerning the ECB's press conferences. Most recently, Edison and Carcel (2020) employ LDA to the Federal Open Market Committee (FOMC) transcripts in order to examine the evolution of topics over time, whereas Hansen et al. (2019) use LDA to identify main topics in the Bank of England's Inflation Report.

## 5.1 Alternative tone based on inflation texts

As explained above, we separate inflation texts from the rest of our textual dataset using unsupervised clustering algorithm, LDA.<sup>16</sup> Our collection of documents consists of all the introductory statements. As we treat each paragraph as a single document, our corpus consists of 3,183 documents and 143,220 words in total.

Number of topics  $K$  needs to be defined for LDA as an input. As the introductory statements are formal and follow a rather fixed structure, it is reasonable to assume that the number of separate topics is relatively small. Thus, for a number of topics  $K$ , we consider values from 2 to 20. We interpret resulting topics with each value of  $K$ , and choose an optimal value based on the interpretability of inflation topics. Within 9 topics, we can identify 2 topics clearly associated with inflation. When the number of topics is increased, we can identify overlapping topics indicating too large a value of  $K$ . In contrast, with smaller values of  $K$ , single topics seem to cover different themes and it is not as straightforward to identify inflation topics. As a result, we choose  $K = 9$  as an optimal number of topics. The list of identified topics is presented in Appendix 3.

After creating a corpus of inflation texts, we calculate an alternative tone with the same procedure as we calculated the general tone. Figure 8 shows the history of the alternative tone index based only on inflation texts in the introductory statements. Comparison of Figures 2 and 8 indicates that, broadly speaking, the evolution of the general tone and the alternative tone have been quite similar even though inflation texts comprise smaller an amount of text compared to all the text in the introductory statements. The share of

---

<sup>16</sup>In each document, a fraction of topics  $\theta$  follows a Dirichlet distribution with a parameter  $\alpha$  and for each topic in a document, a distribution of words follows a Dirichlet distribution with a parameter  $\beta$ . These Dirichlet distributions are used as conjugate priors for a multinomial posterior distribution of topics and conditional multinomial probability of words. Generative process repeats three steps until convergence (Blei et al., 2003):

1. For each topic  $k$ , choose the word distribution  $\gamma_k \sim Dir(\beta)$
2. For each document  $d$ , choose the topic distribution for each topic  $\theta_d \sim Dir(\alpha)$
3. For each  $N$  words  $w_n$ ,
  - (a) Choose a topic  $z_n \sim Multinomial(\theta)$
  - (b) Choose a word  $w_n$  from  $p(w_n|z_n, \beta)$

To approximate a posterior distribution, we follow several other authors studying central bank communication and choose to use a Markov-chain Monte Carlo method collapsed Gibbs sampling (Griffiths & Steyvers, 2004). See e.g. Hansen and McMahon (2019), Kawamura et al. (2019) and Edison and Carcel (2020). When training the model with Gibbs sampling, we use the whole corpus in a learning process.

inflation texts out of all the text is 33%, and the correlation between the two indexes is 0.76. There is some more quarter-to-quarter variation in the inflation tone and it seems that the inflation tone lags the general tone. This has been especially true in the aftermath of the European sovereign debt crisis. While the general tone started to deteriorate at the end of 2013, there is a visible deterioration in the inflation tone not earlier than at the end of 2015, when deflation pressures started to increase.

## 5.2 Results using alternative tone

Estimation results for V-shaped loss functions acquired by using the inflation tone are reported in Table 3. In the case of a symmetric piecewise linear loss function, the estimated slope parameter  $\hat{\delta}$  (1.170) is highly statistically significant and the implicit *de facto* inflation aim is estimated to be 1.84% (see column (1)). Both these parameter estimates are slightly higher than the corresponding estimates based on the general tone with a similar specification.

In column (2) of Table 3, we consider an asymmetric specification. We fix the inflation aim to 1.84% and estimate separate parameters for negative and positive inflation gaps. In this case, the estimated coefficients  $\hat{\delta}_B$  and  $\hat{\delta}_A$  (-1.195 and 0.998, respectively) remain relatively close in absolute terms to the corresponding point estimate  $\hat{\delta}$  in the symmetric loss function (1.170) and the F-test does not reject the hypothesis of symmetry. Hence, the results are qualitatively unchanged compared to those based on the general tone with a fixed target inflation rate of 1.66 (see column (3) of Table 1).

We have also estimated an asymmetric specification without fixing in advance the inflation aim of the ECB. According to column (3) of Table 3, the implicit *de facto* target of the ECB is again 1.7%. The F-test does not reject the hypothesis that the ECB has symmetric preferences with respect to inflation below and above this *de facto* target. Recall that in column (5) of Table 1, we obtained similar results, using the general tone.

In column (4), the *de facto* inflation aim of the ECB is fixed to 2.0%. The estimated slope coefficients  $\hat{\delta}_B$  and  $\hat{\delta}_A$  are of similar size in absolute terms (-0.970 and 1.004), and contrary to the corresponding results in Table 1 (column (8) and, to some extent, column (6)), the F-test does not reject the hypothesis according to

which the ECB has symmetric preferences, even if the *de facto* inflation aim is (assumed to be) 2.0%. The LR test does not reject a *de facto* inflation aim being at 2.0%.

Table 4 summarizes our findings when the alternative tone is used in the estimation of linear-exponential loss functions. The results are again consistent with Table 3, in which we report estimation results based on piecewise linear loss functions. When we estimate all the parameters, the *de facto* inflation target of the ECB is 1.91% and the parameter  $\theta$  is roughly zero, suggesting that the ECB has symmetric preferences (see column (4) in Table 4). When fixing the *de facto* aim of inflation to 1.84%, 1.70% or 2.00%, we also find that the parameter  $\vartheta$  is not significantly different from zero (see columns (1) to (3) of Table 4, respectively). Hence, when using the tone based on inflation text only, we cannot reject (in any of the estimated specifications) the hypothesis according to which the ECB's preferences are symmetric.

Overall, the estimation results based on the general tone and the alternative tone reflecting only inflation texts are qualitatively similar with respect to the level of the *de facto* inflation aim of the ECB. However, the estimations using the general tone support the view according to which ECB policy responses are asymmetric if the inflation aim is set to 2.0%, i.e. to the upper bound of the price stability definition.

## 6 Concluding remarks

In this paper, we have used text analysis methods to assess the characteristics of the ECB's Governing Council's (GC) loss function, and in particular its possible asymmetry and the *de facto* inflation aim. In the estimations, we proxy the ECB's loss by the general tone (net negativity) in the ECB's introductory statements. The tone has then been regressed on the inflation gap, which is constructed by using real time inflation projections of the Eurosystem/ECB staff, available to the GC at the time of the actual policy decisions. We have also allowed for the tone to depend on measures of real activity from the Eurosystem/ECB staff projections. We have repeated our analysis using an alternative proxy for the tone based only on inflation texts.



Overall, our empirical analyses support, even surprisingly well, the earlier findings by Hartmann and Smets (2018), Rostagno et al. (2019) and Paloviita et al. (forthcoming). Most of our results appear to support the notion that the ECB's preferences have been symmetric, with the *de facto* inflation aim clearly below 2% perhaps in the range of 1.6% to 1.7%. If, however, we fix the *de facto* inflation target to 2.0% (the upper bound of the price stability definition), inflation outcomes above 2% have been perceived more costly than below 2%, reflecting an asymmetric loss function. This result holds true in both the piecewise linear and linear-exponential specification of the preferences, but only if we proxy the ECB's loss using the general tone.

The alternative tone based only on a smaller number of words – inflation texts only – shows more quarter-to-quarter variation. The tone based on inflation texts also lags somewhat the general tone. In the monetary policy meetings, the GC evaluates the appropriate monetary policy stance on the basis of a large set of real time economic indicators (not only inflation). Therefore, it is likely that the general tone is more informative on the GC's preferences and its views on both the *de facto* inflation aim and possible asymmetry in monetary policy making.

Both of the two interpretations of the loss function are problematic in the current low rate environment. Asymmetry and a low inflation aim anchor inflation expectations and actual inflation to low levels, generating a low-inflation bias to the policy. This increases the probability of the euro area economy to hit the effective interest rate lower bound and reduces the monetary policy space in the face of future negative shocks. These features of the estimated loss function reflect (with a high likelihood) the ECB's current two-handed definition of price stability and hence raise the question whether and how it should be altered.

Rostagno et al. (2019) provide some narrative and econometric evidence on low inflation bias in the ECB's monetary policy. Their results suggest that while the ECB has tried to keep the headline inflation below the 2% ceiling, it has meant at the same time (in overall) that tighter policy has suppressed the underlying (core) inflation and inflation expectations in the euro area. While this policy has been stabilizing during the period of high inflationary pressures, it has turned out to be costly and destabilizing in the post-financial crisis years when inflationary pressures have been low. While our results do not provide evidence on this

chain of reasoning, our paper provides further evidence that the ECB's has been more averse to high inflation than to low inflation, which is an important part of Rostagno et al. (2019) narrative.

## References

- Armeliu, H., Bertsch, C., Hull, I., & Zhang, X. (2020). Spread the Word: International spillovers from central bank communication. *Journal of International Money and Finance*. doi:<https://doi.org/10.1016/j.jimonfin.2019.102116>
- Baranowski, P., Bennani, H., & Doryń, W. (2020). Do ECB introductory statements help to predict monetary policy: evidence from tone analysis. NBP Working Paper, No. 323.
- Bennani, H., Fanta, N., Gertler, P., & Horvath, R. (2020). Does central bank communication signal future monetary policy in a (post)-crisis era? The case of the ECB. *Journal of International Money and Finance*, 104 (2020) 102167.
- Berger, H., Ehrmann, M., & Fratzscher, M. (2011). Monetary Policy in the Media. *Journal of Money, Credit and Banking*, Vol. 43, No. 4.
- Bholat, D. (2015). Big Data and central banks. *Big Data & Society*. doi:<https://doi.org/10.1177/2053951715579469>
- Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, Vol. 3, 993-1022.
- Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D.-J. (2008). Central Bank Communication and Monetary Policy: a Survey of Theory and Evidence. NBER Working Paper Series, Working Paper 13932.
- Caglayan, M., Jehan, Z., & Mouratidis, K. (2016). Asymmetric monetary policy rules for an open economy: evidence from Canada and the UK. *International Journal of Finance and Economics*. doi:[10.1002/ijfe.1547](https://doi.org/10.1002/ijfe.1547)
- Capistrán, C. (2008). Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious? *Journal of Monetary Economics*, 55, 1415-1427.
- Cassou, S. P., Scott, C. P., & Vázquez, J. (2012). Optimal monetary policy with asymmetric preferences for output. *Economics Letters*, 117, 654-656.
- Christoffel, K., Coenen, G., & Warne, A. (2008). The New Area-Wide Model of the Euro Area: A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis. ECB Working Paper No. 944, October.
- Coibion, O., & Gorodnichenko, Y. (2011). Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation. *American Economic Review*, 101 (1). doi:[10.1257/aer.101.1.341](https://doi.org/10.1257/aer.101.1.341)
- Edison, H., & Carcel, H. (2020). Text data analysis using Latent Dirichlet Allocation: an application to FOMC transcripts. *Applied Economics Letters*. doi:<https://doi.org/10.1080/13504851.2020.1730748>
- Ehrmann, M., & Talmi, J. (2020). Starting from a blank page? Semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, 111, 48-62. doi:<https://doi.org/10.1016/j.jmoneco.2019.01.028>

- Fawcett, N., Koerber, L., Masolo, R., & Waldron, M. (2015). Evaluating UK Point and Density Forecasts from an Estimated DSGE Model: the Role of Off-Model Information Over the Financial Crisis. Bank of England Working Paper, 538.
- Fujiwara, I. (2005). Is the Central Bank's Publication of Economic Forecasts Influential? *Economics Letters*, 89, 255-261.
- Fujiwara, I., & Waki, Y. (2019). Private News and Monetary Policy - Forward Guidance as Bayesian Persuasion. CAMA Working Paper, 91.
- Galí, J. (2015). *Monetary policy, inflation and the business cycle: an introduction*. Princeton University Press.
- Griffiths, T., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101 Suppl 1(1):5228-35. doi:<https://doi.org/10.1073/pnas.0307752101>
- Haldane, A., & McMahon, M. (2018). Central Bank Communications and the General Public. *AEA Papers and Proceedings*, 108, 578-583. doi:<https://doi.org/10.1257/pandp.20181082>
- Hansen, S., & McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*. doi:<https://doi.org/10.1016/j.jinteco.2015.12.008>
- Hansen, S., McMahon, M., & Tong, M. (2019). The long-run information effect of central bank communication. *Journal of Monetary Economics*, 108, 185-202. doi:<https://doi.org/10.1016/j.jmoneco.2019.09.002>
- Hartmann, P., & Smets, F. (2018). The First Twenty Years of the European Central Bank: Monetary Policy. ECB Working Paper No. 2219. Retrieved from <https://ssrn.com/abstract=3309645>
- Hubert, P. (2014). FOMC Forecasts as a Focal Point for Private Expectations. *Journal of Money, Credit and Banking*, 46(7).
- Hubert, P. (2015). *Journal of Money, Credit and Banking*. Do Central Bank Forecasts Influence Private Agents? Forecasting Performance versus Signals, 47(4), 771-789.
- Hubert, P. (2017). Qualitative and Quantitative Central Bank Communication and Inflation Expectations. *Journal of Macroeconomics*, 17(1), 1-41.
- Iversen, J., Laséen, S., H., L., & Söderström, U. (2016). Real-Time Forecasting for Monetary Policy Analysis: the Case of Sveriges Riksbank. *Sveriges Riksbank Working Paper Series*, 318.
- Jones, J. T., Sinclair, T. M., & Stekler, H. O. (forthcoming). A textual analysis of Bank of England growth forecasts. *International Journal of Forecasting*.
- Karagedikli, Ö., & Lees, K. (2007). Do inflation targeting central banks behave asymmetrically? Evidence from Australia and New Zealand. *Reserve Bank of New Zealand Discussion Paper Series*, DP 2004/02.
- Kawamura, K., Kobashi, Y., Shimizu, M., & Ueda, K. (2019). Strategic central bank communication: Discourse analysis of the Bank of Japan's Monthly Report. *Journal of Economic Dynamics and Control*, 230-250. doi:<https://doi.org/10.1016/j.jedc.2018.11.007>

- Kontogeorgos, G., & Lambrias, K. (2019). An Analysis of the Eurosystem/ECB Projections. ECB Working Paper, 2291.
- Korhonen, I., & Newby, E. (2019). Mastering Central Bank Communication Challenges via Twitter. BoF Economics Review.
- Kumar, P., & Jaiswal, U. (2016). A Comparative Study on Sentiment Analysis and Opinion Mining. International Journal of Engineering and Technology.
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual. The Journal Of Finance. doi:<https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. Journal of Accounting Research. doi:<https://doi.org/10.1111/1475-679X.12123>
- Lyziak, T., & Paloviita, M. (2017). Anchoring of Inflation Expectations in the Euro Area: Recent Evidence Based on Survey Data. European Journal of Political Economy, 46, 52-73.
- Lyziak, T., & Paloviita, M. (2018). On the Formation of Inflation Expectations in Turbulent Times: The Case of the Euro Area. Economic Modelling, 72, 132-139.
- Manning, C., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.
- Moreas, R., Valiati, J., & Neto, W. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Systems with Applications, 40(2), 621-633. doi:<https://doi.org/10.1016/j.eswa.2012.07.059>
- Nobay, A. R., & Peel, D. A. (2003). Optimal Discretionary Monetary Policy in a Model of Asymmetric Central Bank Preferences. The Economic Journal, 657-665.
- Orphanides, A., & Wilcox, D. (2002). The Opportunistic Approach to Disinflation. International Finance, 47-71.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). The measurement of meaning. Urbana, IL: University of Illinois Press.
- Paloviita, M., Haavio, M., Jalasjoki, P., & Kilponen, J. (Forthcoming). What Does "Below, But Close To, Two Percent" Mean? International Journal of Central Banking.
- Picault, M., & Renault, T. (2017). Words are not all created equal: A new measure of ECB communication. Journal of International Money and Finance, 79, 136-156.
- Potter, S. (2011). The Failure to Forecast the Great Recession. Retrieved from <https://libertystreeteconomics.newyorkfed.org/>
- Romer, C., & Romer, D. (2000). Federal Reserve information and the behavior of interest rates. American Economic Review, 90 (3), 429-457.
- Romer, C., & Romer, D. (2004). A New Measure of Monetary Shocks: Derivation and Implications. American Economic Review, 94 (4), 1055-1084. doi:10.1257/0002828042002651

- Rostagno, M., Altavilla, C., Carboni, G., Lemke, W., Motto, R., Guilhem, A., & Yiangou, J. (2019). A Tale of Two Decades: The ECB's Monetary Policy at 20. ECB Working Paper No. 2346. Retrieved from <https://ssrn.com/abstract=3515400>
- Ruge-Murcia, F. J. (2007). Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics & Control*, 2599-2636.
- Schmeling, M., & Wagner, C. (2019). Does Central Bank Tone Move Asset Prices? CEPR Discussion Paper No. DP13490. Retrieved from <https://ssrn.com/abstract=3324225>
- Shapiro, A., & Wilson, D. (2019). Taking the Fed at its Word: Direct Estimation of Central Bank Objectives using Text Analytics. Federal Reserve Bank of San Francisco Working Paper 2019-02. doi:10.24148/wp2019-02
- Sharma, A., & Dey, S. (2012). A comparative study of feature selection and machine learning techniques for sentiment analysis. *RACS '12: Proceedings of the 2012 ACM Research in Applied Computation Symposium*, (pp. 1-7). doi:<https://doi.org/10.1145/2401603.2401605>
- Stockton, D. (2012). Review of the Monetary Policy Committee's Forecasting Capability. Report Presented to the Court of the Bank of England.
- Surico, P. (2003). Asymmetric Reaction Functions for the Euro Area. *Oxford Review of Economic Policy*, 19(1), 44-57.
- Surico, P. (2007). The Fed's monetary policy rule and U.S. inflation: The case of asymmetric preferences. *Journal of Economic Dynamics & Control*, 31, 305-324.
- Tobback, E., Nardelli, S., & Martens, D. (2017). Between hawks and doves: measuring central bank communication. ECB Working Paper Series, European Central Bank (ECB). doi:10.2866/977130
- Walsh, C. (2003). Speed Limit Policies: The Output Gap and Optimal Monetary Policy. *American Economic Review*, 265-278. doi:10.1257/000282803321455278
- Woodford, M. (2003). *Interest and Prices*. Princeton University Press.

**Table 1. Estimated V-shaped loss function, general tone**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Symmetric V	Symmetric V	Asymmetric V	Asymmetric V	Asymmetric V	Asymmetric V	Asymmetric V	Asymmetric V
$\alpha$	-1.0552***	-1.5695***	-1.0909***	-1.6508***	-1.0776***	-0.8711***	-1.4469***	-1.4895***
$\delta$	1.1035***	0.7475***						
$\delta_B (D = 0)$			-1.0392***	-0.5477*	-0.9977***	-0.6342**	-0.2216	-0.3086
$\delta_A (D = 1)$			1.2525***	1.0967***	1.3002***	1.5286***	2.2903***	1.6002***
$\Delta y$		-1.2599***		-1.2286**			-1.3024***	-1.3154***
$\Delta y^2$		-0.1161**		-0.0970*			-0.1054**	-0.1097**
Observations	81	81			81	81	81	81
Adjusted R <sup>2</sup>	0.1	0.50	0.09	0.51	0.09	0.08	0.55	0.58
$\bar{\pi}^+$	1.66	1.66			1.70		2.3	
$\pi_{fixed}^+$			1.66	1.66		2		2
F-test symmetry p-value	-	-	0.6032	0.082*	0.4684	0.0897*	0.0001***	0.0014***
Likelihood-ratio test p-value						0.1204		0.2218
95% C.I.	[1.30, 1.95]	[0.11, 2.05]						

95% confidence interval is obtained using bootstrap procedure

Likelihood – ratio test is performed between restricted model ( $\pi_{fixed}^+$ ) and unrestricted model ( $\bar{\pi}^+$ )

\*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01

**Table 2. Estimated Lindex loss function, general tone**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
constant	-5.6021	-4.2452	-1.3292*	-3.2975	-2.8853***	-3.1913**	-1.8476***	-3.0514
$\alpha$	0.449	0.5263	1.1661**	0.612	0.8298***	0.7561**	1.4624***	0.7873
$\varphi_1$					-0.1149**	-0.1139***	-0.1184**	-0.1144**
$\varphi_2$					-1.2874***	-1.2854***	-1.2959***	-1.2863***
Observations	81	81	81	81	81	81	81	81
$\bar{\pi}^+$				1.74***				1.68***
$\pi_{fixed}^+$	1.66	1.70	2		1.66	1.70	2	
Likelihood-ratio test p-value	0.786	0.888	0.446		0.948	0.930	0.261	

Likelihood – ratio test is performed between restricted model ( $\pi_{fixed}^+$ ) and unrestricted model ( $\bar{\pi}^+$ )

\*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01

**Table 3. Estimated V-shaped loss function, alternative tone**

	(1)	(2)	(3)	(4)
	Symmetric V	Asymmetric V	Asymmetric V	Asymmetric V
$\alpha$	-0.2204	-0.1838	-0.2238	-0.0897
$\delta$	1.1696***			
$\delta_B (D = 0)$		-1.1946***	-1.3785***	-0.9703***
$\delta_A (D = 1)$		0.9977***	0.9014**	1.0004*
Observations	81		81	81
Adjusted R <sup>2</sup>	0.15	0.09	0.15	0.11
$\widehat{\pi}$	1.84		1.70	
$\pi_{fixed}^*$		1.84		2
F-test symmetry p-value	-	0.6547	0.2385	0.9527
Likelihood-ratio test				0.978

*Likelihood – ratio test is performed between restricted model ( $\pi_{fixed}^*$ ) and unrestricted model ( $\widehat{\pi}$ )*

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 4. Estimated Linex loss function, alternative tone**

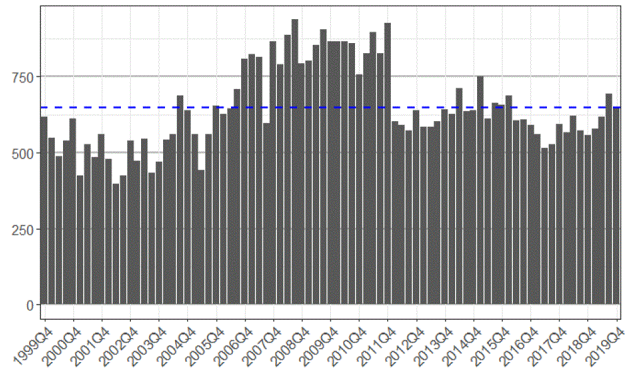
	(1)	(2)	(3)	(4)
$\alpha$	0.1449	0.1407	0.1471	0.1465
$\theta$	-0.1391	-0.3992	0.1346	-0.022
$\gamma$	1.0892**	1.0874**	1.0837**	1.0871**
Observations	81	81	81	81
$\widehat{\pi}$				1.91
$\pi_{fixed}^*$	1.84	1.70	2	
Likelihood-ratio test p-val	0.786	0.888	0.446	

*Likelihood – ratio test is performed between restricted model ( $\pi_{fixed}^*$ ) and unrestricted model ( $\widehat{\pi}$ )*

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$



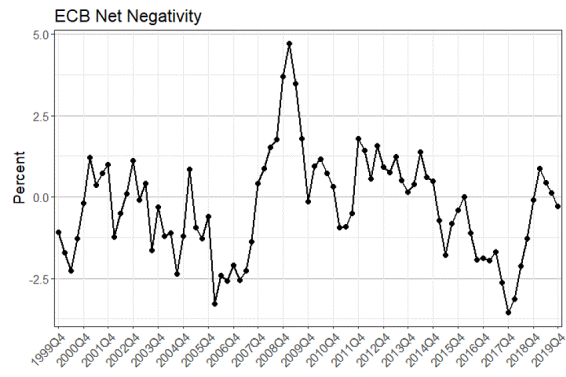
**Figure 1. Quarterly average of the number of words in the ECB's introductory statements 1999Q4-2019Q4**



**Note:** The dashed blue line corresponds to the overall mean.

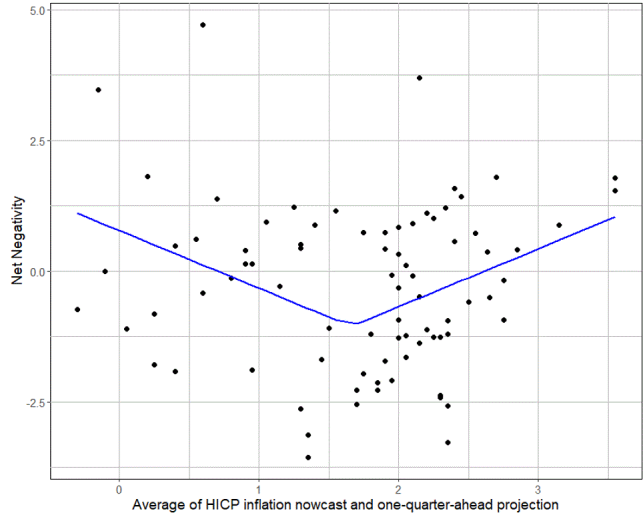
**Sources:** ECB and authors' own calculations.

**Figure 2. General tone**



**Sources:** ECB and authors' own calculations.

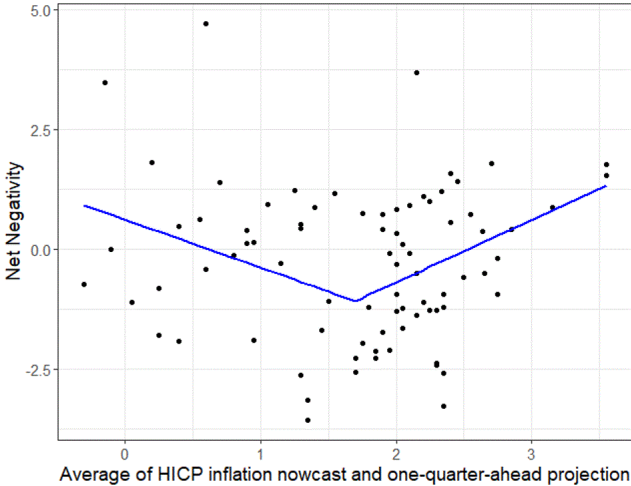
**Figure 3. Estimated symmetric loss function**



**Note:** The estimated loss function of Figure 3 corresponds to column (1) in Table 1.

**Sources:** ECB and authors' own calculations.

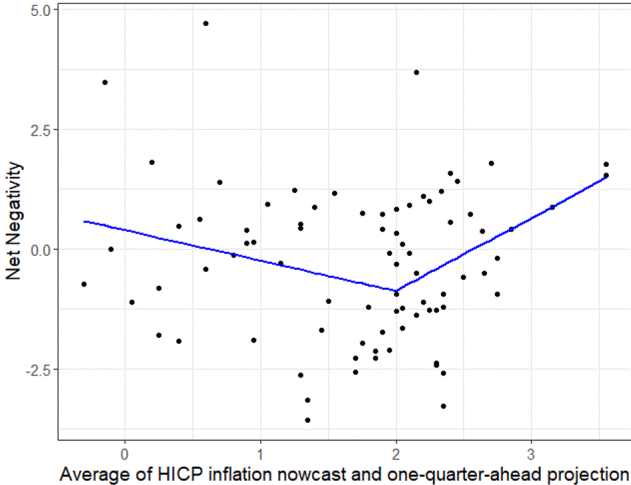
**Figure 4. Estimated asymmetric loss function**



**Note:** The estimated loss function of Figure 4 corresponds to column (5) in Table 1.

**Sources:** ECB and authors' own calculations.

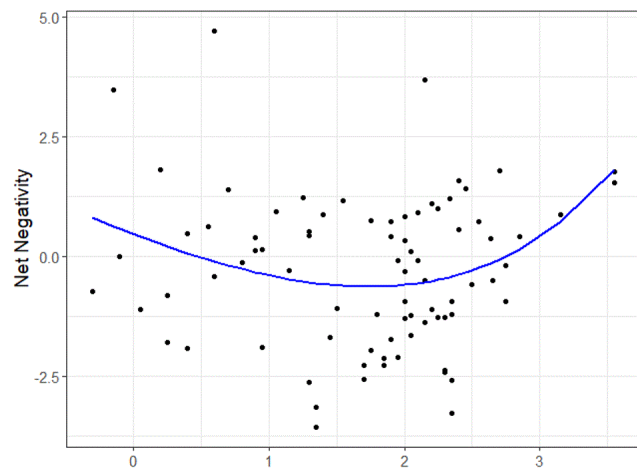
Figure 5. Estimated asymmetric loss function, fixed  $\pi^*=2.0$



**Note:** The estimated loss function of Figure 5 corresponds to column (6) in Table 1.

**Sources:** ECB and authors' own calculations.

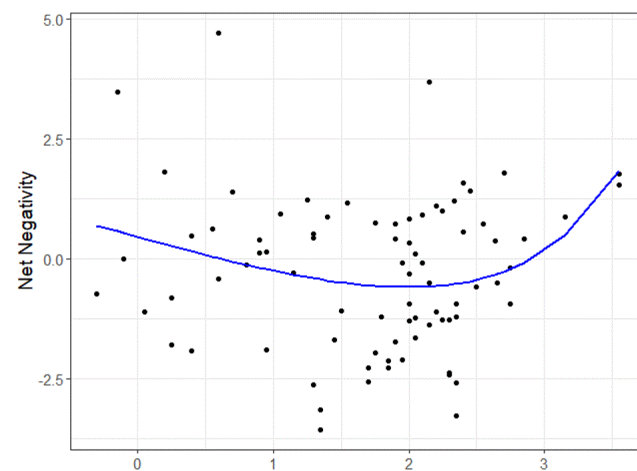
**Figure 6. Estimated Linex loss function**



**Note:** The estimated loss function of Figure 6 corresponds to column (4) in Table 2.

**Sources:** ECB and authors' own calculations.

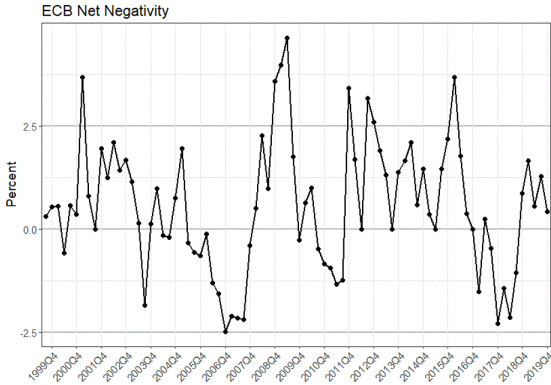
**Figure 7. Estimated Linex loss function, fixed  $\pi^*=2.0$**



**Note:** The estimated loss function of Figure 7 corresponds to column (3) in Table 2.

**Sources:** ECB and authors' own calculations.

Figure 8. Alternative tone based on inflation texts



Sources: ECB and authors' own calculations.

## APPENDIX 1

### Example of scored text, the ECB's introductory statement

December 12, 2019

[...]

“The incoming data since our last meeting are in line with our baseline scenario of ongoing, but moderate, growth of the euro area economy. In particular, the **weakness** in the manufacturing sector remains a **drag** on euro area growth momentum. However, ongoing, albeit decelerating, employment growth and increasing wages continue to support the resilience of the euro area economy. While inflation developments remain subdued overall, there are some signs of a moderate increase in underlying inflation in line with expectations.

The unfolding monetary policy measures are underpinning **favourable** financing conditions for all sectors of the economy. In particular, **easier** borrowing conditions for firms and households are supporting consumer spending and business investment. This will sustain the euro area expansion, the build-up of domestic price pressures and, thus, the **robust** convergence of inflation to our medium-term aim.”

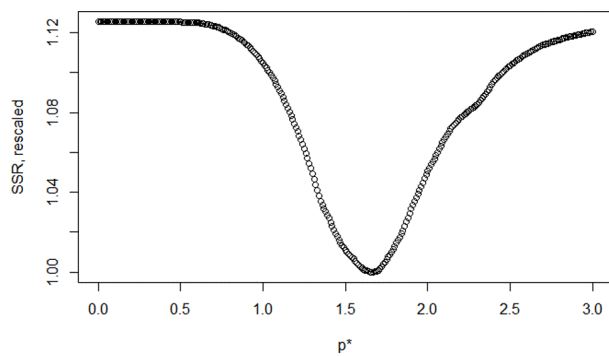
[...]

“Euro area annual HICP inflation increased to 1.3% in December 2019, from 1.0% in November, reflecting mainly higher energy price inflation. On the basis of current futures prices for oil, headline inflation is likely to hover around current levels in the coming months. While indicators of inflation expectations remain at low levels, recently they have either stabilised or ticked up slightly. Measures of underlying inflation have remained generally muted, although there are further indications of a moderate increase in line with previous expectations. While labour cost pressures have **strengthened** amid tighter labour markets, the **weaker** growth momentum is **delaying** their pass-through to inflation. Over the medium term, inflation is expected to increase, supported by our monetary policy measures, the ongoing economic expansion and solid wage growth.”

[...]

## APPENDIX 2

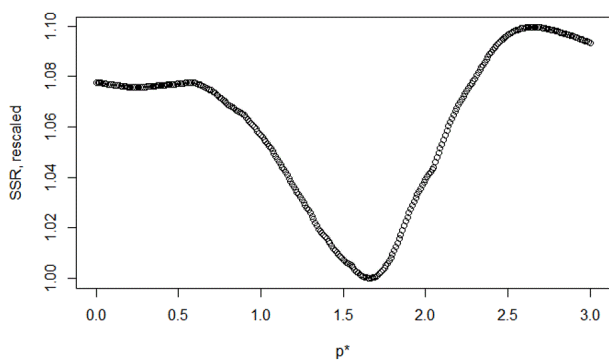
Figure A2.1. Sums of squared residuals



**Note:** The sums of squared residuals (SSR) for column (1) of Table 1. The SSRs are rescaled so that a value of 1 corresponds to the minimum.

**Source:** Authors' own calculations.

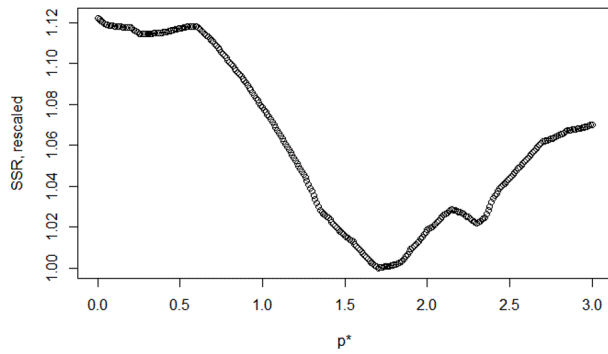
Figure A2.2. Sums of squared residuals



**Note:** The sums of squared residuals (SSR) for column (2) of Table 1. The SSRs are rescaled so that a value of 1 corresponds to the minimum.

**Sources:** Authors' own calculations.

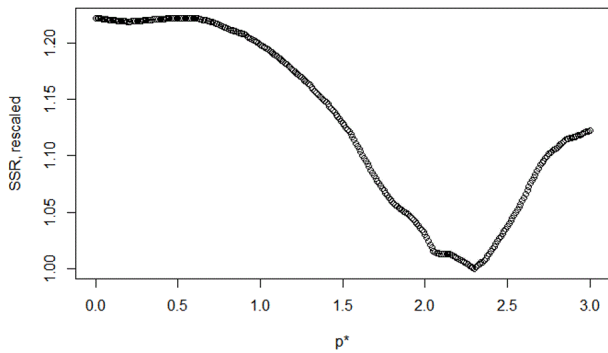
**Figure A2.3. Sums of squared residuals**



**Note:** The sums of squared residuals (SSR) for column (5) of Table 1. The SSRs are rescaled so that a value of 1 corresponds to the minimum.

**Source:** Authors' own calculations.

**Figure A2.4. Sums of squared residuals**



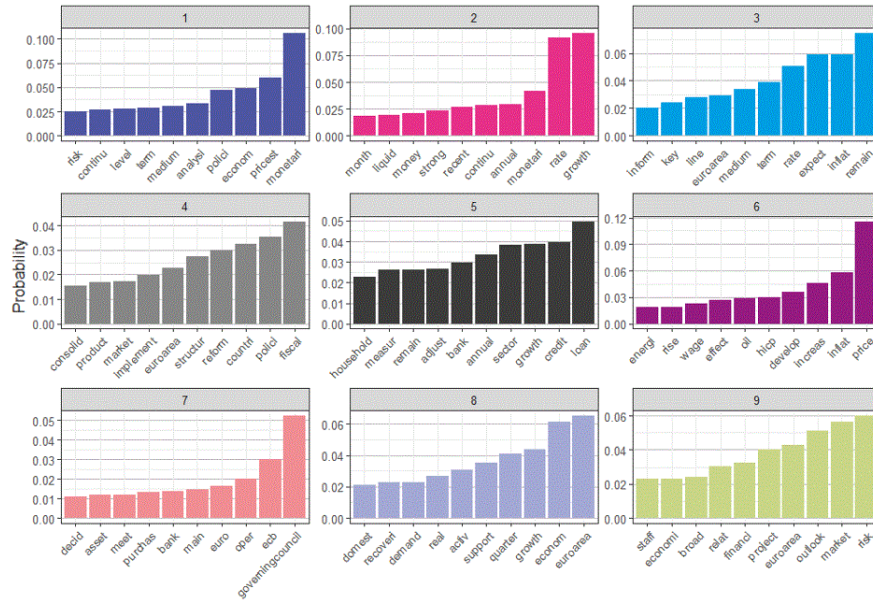
**Note:** The sums of squared residuals (SSR) for column (7) of Table 1. The SSRs are rescaled so that a value of 1 corresponds to the minimum.

**Sources:** Authors' own calculations.



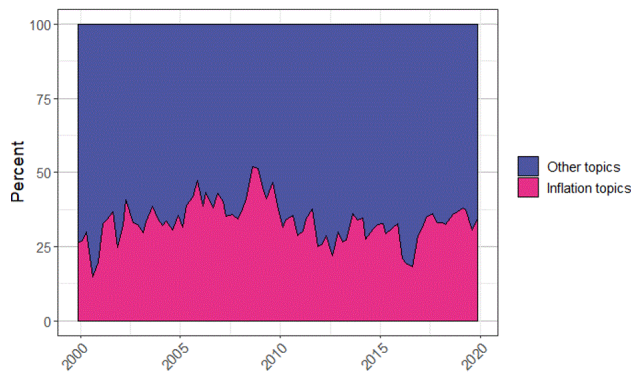
## APPENDIX 3

Figure A3.1. Topics discovered by LDA



Sources: ECB and authors' own calculations.

Figure A3.2. Fractions of topics in the ECB's introductory statements



Sources: ECB and authors' own calculations.

# Bank of Finland Research Discussion Papers 2020

ISSN 1456-6184, online

- 1/2020 Masashige Hamano – Francesco Pappada  
**Firm turnover in the export market and the case for fixed exchange rate regime**  
ISBN 978-952-323-309-6, online
- 2/2020 Gonçalo Faria – Fabio Verona  
**Frequency-domain information for active portfolio management**  
ISBN 978-952-323-310-2, online
- 3/2020 Tomi Kortela – Jaakko Nelimarkka  
**The effects of conventional and unconventional monetary policy: identification through the yield curve**  
ISBN 978-952-323-311-9, online
- 4/2020 Manuel M. F. Martins – Fabio Verona  
**Forecasting inflation with the New Keynesian Phillips curve: Frequency matters**  
ISBN 978-952-323-322-5, online
- 5/2020 Gene Ambrocio  
**Inflationary household uncertainty shocks**  
ISBN 978-952-323-324-9, online
- 6/2020 Gonçalo Faria – Fabio Verona  
**Time-frequency forecast of the equity premium**  
ISBN 978-952-323-325-6, online
- 7/2020 Eleonora Granziera – Markus Sihvonen  
**Bonds, Currencies and Expectational Errors**  
ISBN 978-952-323-326-3, online
- 8/2020 Satu Nurmi – Juuso Vanhala – Matti Virén  
**The life and death of zombies – evidence from government subsidies to firms**  
ISBN 978-952-323-328-7, online
- 9/2020 Aino Silvo – Fabio Verona  
**The Aino 3.0 model**  
ISBN 978-952-323-329-4, online
- 10/2020 Gene Ambrocio – Iftekhar Hasan – Esa Jokivuolle – Kim Ristolainen  
**Are bank capital requirements optimally set? Evidence from researchers' views**  
ISBN 978-952-323-330-0, online
- 11/2020 Giovanni Caggiano – Efrem Castelnuovo – Richard Kima  
**The global effects of Covid-19-induced uncertainty**  
ISBN 978-952-323-332-4, online
- 12/2020 Maritta Paloviita – Markus Haavio – Pirkka Jalasjoki – Juha Kilponen – Ilona Vänni  
**Reading between the lines – Using text analysis to estimate the loss function of the ECB**  
ISBN 978-952-323-337-9, online