



IFC-Bank Indonesia Satellite Seminar on “*Big Data*” at the ISI Regional Statistics Conference
2017

Bali, Indonesia, 21 March 2017

Between hawks and doves: measuring Central Bank Communication¹

Stefano Nardelli, European Central Bank,
David Martens and Ellen Tobback, University of Antwerp

¹ This paper was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.

Between Hawks and Doves: Measuring Central Bank Communication

Stefano Nardelli¹, David Martens² and Ellen Tobback³

Abstract

Media scrutinise the ECB's communication very attentively to extract information on likely future moves of monetary policy rates, in particular after each press conference following monetary policy meetings. Assessing media's perception requires the translation of words into a quantitative indicator. In this paper, we propose the Hawkish-Dovish (HD) indicator which is computed out of a bulk of above 9,000 media reports on the ECB press conference (i.e. newspaper articles, newswires, etc.) and translates on a numerical scale the degree of "hawkishness" or "dovishness" of the ECB press conference tone perceived by media. We compare two different methods to calculate the indicator: one is based on the semantic orientation, while the other is constructed on a text classification with a Support Vector Machines classification model. We show that the latter method tends to provide more stable and accurate measurements of the perception on a labelled test set. Furthermore, we demonstrate the potential use of this indicator with some applications. We analyse the correlations with a set of interest rates and demonstrate the indicator's ability to anticipate monetary policy moves. Additionally, we use the Latent Dirichlet Allocation (LDA) algorithm to detect the dominant topics in the articles on ECB monetary policy decisions and conclude that the media's focus has shifted from the classic interest rate movements towards the non-standard monetary policy instruments. These findings provide decisive evidence in favour of using an advanced text mining classification model to measure more accurately how media perceive ECB communication.

Keywords: Central bank policies, monetary policy, communication, big data, data mining, quantitative methods.

JEL classification: C02, C63, E52, E58

¹ This paper was prepared for the IFC-Bank Indonesia Satellite Seminar on "Big Data" at the ISI Regional Statistics Conference 2017 in Bali on 21 March 2017. The authors are grateful to participants to the conference and to all preliminary seminars for their useful comments and fruitful discussions on the paper. The views expressed here do not represent those of the European Central Bank.

Correspondence: Stefano Nardelli; European Central Bank; E-mail: stefano.nardelli@ecb.europa.eu

² University of Antwerp, Belgium; Email: david.martens@uantwerpen.be

³ University of Antwerp, Belgium; Email: ellen.tobback@uantwerpen.be

Contents

Between Hawks and Doves: Measuring Central Bank Communication..... 1

1. Introduction..... 3

2. The HD index..... 4

3. Data and results 5

4. Validation criteria 7

 Performance analysis..... 8

 Correlation analysis 8

 Topic classification..... 9

5. Conclusions..... 11

References..... 12

1. Introduction

During the financial crisis, monetary policy interest rates reached levels close to zero in most major central banks. As a result, communication has increasingly qualified as a powerful instrument able to influence financial market developments and drive expectations. Forward guidance, i.e. a form of verbal commitment to keep interest rates at a certain level for a period of time, possibly conditional to certain economic developments, became a genuine monetary policy instrument for many major central banks in addition to the standard toolkit of interest rates and refinancing instruments for the banking sector. Mirroring the increased role of communication in central banks, economic research has started to focus on how central bank communication adds information to that which is already contained in macroeconomic variables and how it may reveal policy makers individual preference functions. The overall objective is to enhance predictability of monetary policy decisions. The traditional approach focuses mainly on how information events affect both financial market developments and expectations of future policy moves.

Media usually analyse central banks' communication with a view to extract information about likely future moves of monetary policy rates (Hayo and Neuenkirch, 2015). Unlike market indicators, extracting relevant information on monetary policy from media reports is not straightforward and requires the translation of words into a numerical representation. Ideally, the resulting quantitative indicator should represent the media's perception of the central banks message and should be able to indicate whether expectations are more on the tightening side (hawkish perception) or rather on the loosening side (dovish perception).

In recent years, a number of studies have developed quantitative tools to measure or, at least, to express on a numerical scale the information contained in official central banks' statements (e.g. KOF Swiss Economic Institute, 2007; Lucca and Trebbi, 2009; Hansen and McMahon, 2016; and Nechio and Regan, 2016). In other words, this approach can be characterised as an attempt to quantify information in communication, which is qualitative by nature. In this paper, we present an index measuring how the official monetary policy communications of the European Central Bank (ECB) are interpreted by the press, using data mining techniques. Concretely, we want to distinguish articles that perceive the ECB's statement as predominantly hawkish, i.e. pointing towards likely future policy rate increases, from those perceiving statements as dovish, i.e. hinting at declines or no increases in policy rates. We use two different techniques normally used in text analysis to measure the tone expressed in an article published after the press conference: Semantic Orientation (SO) and text classification using Support Vector Machines (SVM). The former measures how often the ECB is mentioned in a news article together with respectively hawkish and dovish words, the latter uses a classification model to predict the tone of a news article. We apply both methods to a data set of approximately 9,000 articles from the Dow Jones's Factiva global news database (which includes several million of items from nearly 33,000 sources), published between January 1999 and March 2016, in order to create the HD index (after the initials of Hawkish and Dovish). This indicator represents a sort of average tone between two hawkish and dovish extremes of the ECB communication as perceived by the media after each press conference.

This paper aims at offering an original contribution on various dimensions. First, it proposes a quantitative indicator computed from text sources, showing that data or text mining techniques can bring value in the process of monetary policy decision making (Section 2 and 3). Secondly, it compares two alternative methodologies and concludes in favour of the use of SVM (Section 4). Finally, it shows that one of the advantages of the SVM classification model there is its flexibility and the possibility to analyse in details terms most frequently used by the written media to describe monetary policy decisions (Section 4.3).

2. The HD index

The HD index was computed using two different methodologies: the first based on semantic orientation and the other on a support vector machine (SVM).

The first approach followed the methodology originally proposed by Lucca and Trebbi (2009), who analysed how Federal Open Market Committee (FOMC) monetary policy statements were reflected in press reports. Semantic orientation is a concept from computational linguistics and defines the position of a word or string of words between two opposite concepts. Turney (2002) has used Semantic Orientation (SO) to classify reviews as positive or negative. In practice, the SO score is defined as the difference between the strength of its association with a set of words associate to two opposite concepts such as "hawkish" and "dovish" in a subset R of a corpus of texts, i.e.:

$$\begin{aligned} SO &= \log\left(\frac{\Pr(\text{Hawkish} \ \& \ R)}{\Pr(\text{Hawkish}) \times \Pr(R)}\right) - \log\left(\frac{\Pr(\text{Dovish} \ \& \ R)}{\Pr(\text{Dovish}) \times \Pr(R)}\right) \\ &= \log\left(\frac{\Pr(\text{Dovish})}{\Pr(\text{Hawkish})}\right) + \log\left(\frac{\Pr(\text{Hawkish} \ \& \ R)}{\Pr(\text{Dovish} \ \& \ R)}\right) \end{aligned}$$

In this way, the values taken on a numeric scale by the SO reflects the relative frequency of two opposite concepts, in this case "hawkish" and "dovish".

The HD index based on this method is computed by counting the co-occurrences of strings with words and expressions that are normally associated with these extreme concepts on a set of texts reporting on ECB monetary policy decisions. Texts are extracted from a corpus of media reports (Dow Jones's Factiva). The formula used is the following:

$$HD_t = \frac{\sum_{s(t)} (-1) \times I[s(t), R, D] + \sum_{s(t)} (+1) \times I[s(t), R, H]}{\sum_{s(t)} I[s(t), R, D] + \sum_{s(t)} I[s(t), R, H]}$$

with $I(\)$ being the indicator function that counts the co-occurrences of a word $s(t)$ pre-classified as dovish (D) or hawkish (H) in a given set of articles referring to a time t extracted from a corpus through a filter $R = \{\text{"European Central Bank" or "ECB" or "Mario Draghi" etc.}\}$. By construction, the HD index takes values in the interval $[-1; +1]$, where -1 indicate a maximum degree of dovishness and $+1$ the opposite.

The alternative methodology applies text mining techniques, i.e. a Support Vector Machines (SVM) classification model following Provost and Fawcett (2013). This technique automatically looks for patterns in text documents to select the words with the highest discriminative power. The output is a linear model in which

each word is assigned a weight in favour of either class +1 (i.e. hawkish) or -1 (i.e. dovish). A clear advantage is that this algorithm looks at every document as a whole and therefore tends to overcome the limitations of a predefined set of keywords as in the previous method.

To initialise the algorithm, a so-called *training set* was formed by selecting approximately 550 articles and pre-classifying them as “hawkish” or “dovish”. These articles were randomly selected from the available corpus. However, articles of uncertain classification were excluded not to introduce any bias and to enhance prediction accuracy. All articles were transformed into a ‘bag-of-words’ vector, i.e. $[t_0 t_1 \dots t_j \dots t_n]$ containing all n unique words present in the training set, with t_j being the occurrences of word j in the article. From such vector a term-frequency matrix $tf(m, n)$ is built – with m being the number of articles and n the total number of words and – in which each cell (i, j) indicates the number of times a word j occurs in article i .

In order to reduce the relative weight of words occurring very frequently in the training set of articles, each term count is multiplied by the inverse document frequency (*idf*), which measures the frequency of a term across all documents (Weiss et al., 2010), i.e.:

$$idf(t, m) = \log \frac{\text{Number of articles in the training set } (m)}{\text{Number of articles in the training set where term } t \text{ occurs}}$$

The resulting *tf-idf* matrix is used as input to the SVM algorithm searching for the “decision boundary” maximising the margin between the two classes, i.e. “dovish” and “hawkish” in this case.

Linear SVM tries to solve the following optimization problem (Fan et al., 2008):

$$\min_w \frac{1}{2} w^T w + C \sum_i \max(1 - y_i w^T x_i, 0)^2$$

with w being the vector of the weights in the model, x_i and y_i representing respectively the input vector and the label of the i -th observation, while C is a cost parameter defined exogenously.

Articles are classified based on the following linear model:

$$f(x_i) = w_0 + w_1 x_{i1} + w_2 x_{i2} + \dots + w_j x_{ij} + \dots + w_n x_{in}$$

in which the weights w_j are estimated from the optimisation problem and x_{ij} is the occurrence (frequency) of the j -th unique terms of the training set in the i -th article. The sign of the resulting decision value $f(x_i)$ is the predicted class the article belongs to (i.e. hawkish or dovish), whereas the value of $f(x_i)$ approximates the article’s degree of hawkishness or dovishness. In this way, the larger the decision value is, the more certain the classifier is about the chosen class. Document of uncertain classification would therefore tend to have a decision value around 0, which can be interpreted as neutral tone in this context.

3. Data and results

To compute our HD-index, we used articles extracted from Dow Jones Factiva. The selection was restricted to articles mentioning “ECB” or “European Central Bank” or the name of its President and limited to categories ‘Major News and Business

Publications: Europe' and 'Major News and Business Publications: US'. Only articles in English were included in the sample. Because the focus was the ECB press conference, a three-day window around the press conference (i.e. the day before, the very day and the day after) was used to extract articles. Using these criteria, we formed a corpus of slightly less than 9,000 articles, published between January 1999 and the latest press conference, covering therefore the whole history of the euro (in this paper, however, we included data until January 2016).

Figure 1 shows the news sources included in our data set and the respective proportion of articles in the corpus.

Sources and respective proportion of articles in the corpus

Figure 1

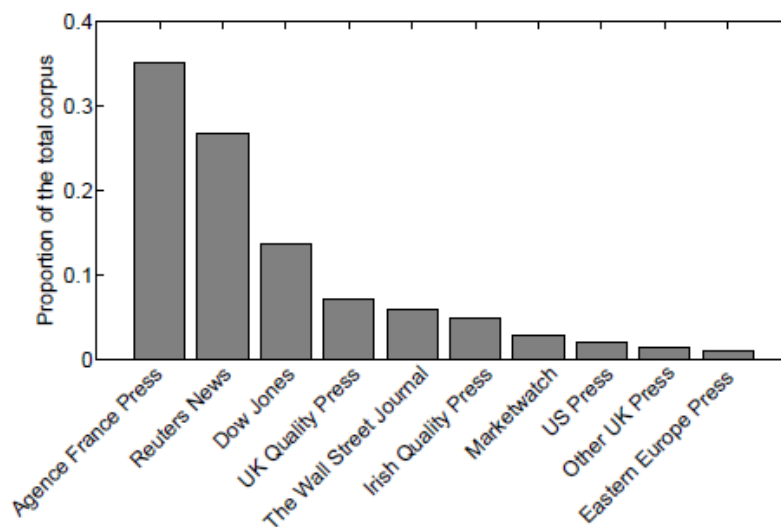


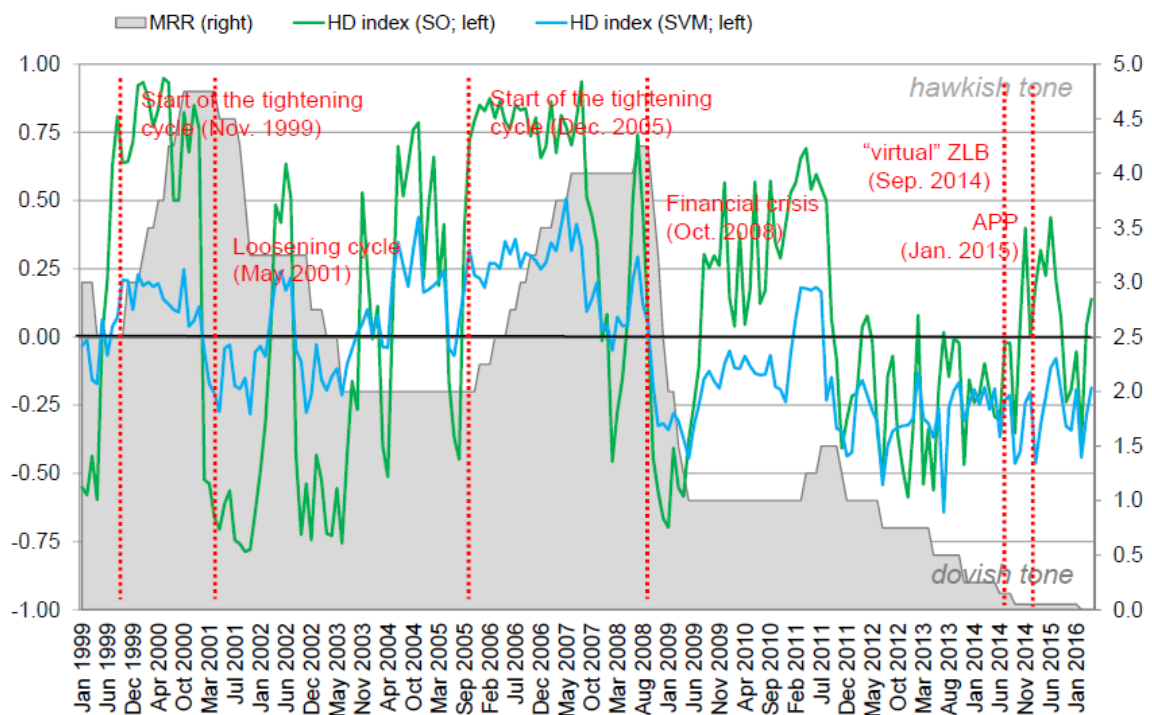
Figure 2 shows the index computed using the two methodologies presented in Section 2 for the whole sample. The two series are contrasted with the evolution of the ECB monetary policy rate (namely, the interest rate on main refinancing operations) to contextualize them. Some relevant episodes are also indicated. Overall, the HD index generally appears to anticipate turning points in the ECB monetary policy stance and to be consistent with the different monetary policy cycles with a few exceptions. In other words, changes in the tone of communication that anticipate official interest rate movements generally appear to be understood appropriately. Specifically, the HD index rose prior to the two tightening monetary policy cycles started at the end of 1999 and at the end of 2005, but also HD declined ahead of the loosening cycles started in mid-2001 and the one caused by the deterioration of the financial crisis following Lehman's collapse in autumn 2008 (in particular the SVM index moved from a very hawkish to a neutral/dovish stance already in late 2007). During the crisis period the SVM HD index has consistently hovered around negative values thus indicating a prevailing dovish tone perceived in the ECB's communication. The decline to relatively low values in 2011 anticipated the marked loosening in monetary policy conditions taking the form of unprecedented standard and non-standard measures in order to bring inflation in line with its objective. While relatively low values were reached in the OMT phase, interestingly the HD index has become relatively less dovish in the period following

the announcement of the extended APP, probably reflecting the belief that the ECB loosening policy reached a bottom or, at least, that what was done was sufficient.

As regards differences in results between the two methodologies, although they broadly show a similar dynamics (the correlation between the two series is 0.78), the SO method tends to produce more marked volatile results and occasionally inconsistent results (for instance, towards the end of 2014).

HD index computed with SO and SVM and ECB official interest rate

Figure 2



4. Validation criteria

Whereas the SVM methodology overcomes some obvious limitations of the SO methodology (the preselection of a fixed set of words in the first place), there is no obvious validation method to compare the results obtained under the two methods. In this section we propose on three different criteria:

1. we measure how well both methods can predict whether the tone of an article is predominantly hawkish or dovish;
2. we analyse in a deeper fashion the link of the HD index with some interest rates through correlations; and
3. we evaluate the indicators' performance qualitatively by extracting the most important topics mentioned in the news articles to investigate if they can be linked with various phases of monetary policy as identified by the HD index.

Performance analysis

The classifications (scores) obtained from the two methodologies are compared with our own manual classifications using two performance metrics: Area under the receiver operating curve (AUC) and accuracy. AUC is a standard evaluation metric for classification models that represents a model's discriminative power by measuring to what extent positively labelled observations (hawkish) are ranked higher than negatively (dovish) labelled observations (Fawcett 2006). Unlike accuracy, which represents the percentage of correctly classified observations, AUC is able to deal with unbalanced distributions. Results are shown in Table 1.

Method	AUC	Accuracy
SO	69.24%	62.00%
SVM	98.55%	92.00%

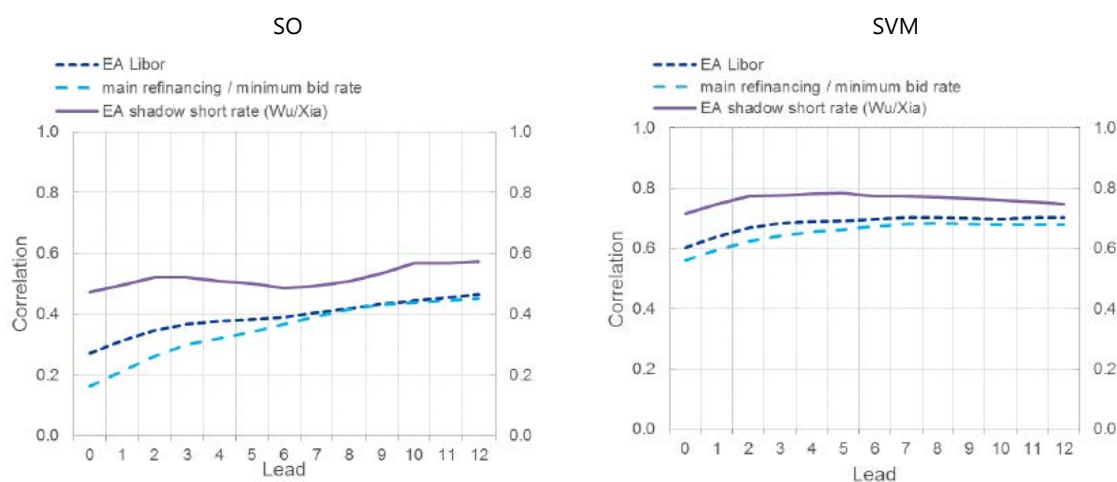
The SVM methodology outperforms SO when classifying the test articles as hawkish or dovish. We find a difference in accuracy of 30 percentage points in favour of SVM. The more advanced SVM-classification model is able to look at the broader context of an article, which results in a dovish classification for the example article.

Correlation analysis

The visual inspection in Section 3 is substantiated here with the analysis of the correlation of the HD index and the ECB main refinancing rate. In addition, correlations were also computed with the Wu-Xia shadow rates to take into account the actual stance when official interest rates reached the zero lower bound and with the 12-month euro LIBOR rate, a money market interest rate reflecting expectations on official rates over the one-year horizon.

Figure 3 shows correlograms from zero to twelve meetings ahead and can be interpreted as an indication on the lead properties of the HD index on actual monetary policy moves. It is important to stress that evidence of correlation should be interpreted here as a positive indication about the ability of the press to interpret correctly communication on future monetary policy stance (and of the index to represent significantly such interpretation).

Positive correlation actually exists for the HD index and tends to be higher for the SVM methodology. The HD index tends to lead actual interest rates moves after 6 to 7 Governing Council meetings and then levels off for the SVM (8 to 9 meetings for SO). Furthermore, a relatively higher correlation (reaching a 0.8 peak) exists with the Wu-Xia shadow rates with a lead of 4 meetings in the case of SVM; also in this case SO tends to have a weaker correlation of only 0.6 and a lead of 11 meetings. A similar picture emerges from the correlation with the 12-month euro LIBOR rate with a lead of approximately 7 meetings and a peak correlation of about 0.7 for HD SVM and 0.45 for HD SO.



Topic classification

In this case, a topic classification model was applied to the dataset divided into 13 periods selected on the basis of official interest rate developments (i.e. rising, stable, declining) and consistently with the values of the HD index. The idea behind is to check whether an unsupervised algorithm is able to identify topics that can be reasonably associate with the various phases of ECB monetary policy. If that is the case, one can conclude that communication successfully conveyed (and reporters were able to relate) the most relevant elements of the monetary policy discourse.

Topics are derived using a Latent Dirichlet Allocation (LDA) as in Blei et al. (2003). This method assumes that each document can be represented by a mixture of topics, where the topic distribution is assumed to follow a Dirichlet prior. Every topic has a probability of generating a set of words, which makes it possible to characterise the topic on the base of the related words. In other words, a topic is defined by a cluster of words, which tend to appear together in a document or corpus and are identified by the LDA algorithm.

The topic subjects and a selection of the top 10 words for these topics can be found in Table 2, whereas Figure 4 shows the three most dominant topics in each period and contrasted with the HD index.

Topics identified by the LDA algorithm tend to be consistent with the tightening cycles (and the hawkish peaks of the HD index). Between May and July 2002 and between June 2004 and April 2005 both topic A – characterised by words as “rise”, “hike”, “increase” as well as “price” associate to a tightening monetary policy – and topic C – characterised by words “growth”, “inflat[ion]”, “price” all pointing to a strengthening of economic cycle – tend to dominate. In tightening and easing cycles, dominant topics are those related to the interest rate movements, i.e. topic A or B. However, topic B (pointing to rate cut) tends also to appear in the same phases, possibly indicating that a level of uncertainty or disagreement not only in periods of stable monetary policy rates but also in those with interest rates increasing. Interestingly, Figure 4 confirms also the evolution in

topics from those predominantly related to interest rate decisions (A and B) or price/growth expectations (C) to crisis (E) and asset/bond purchases (F) after the onset of the financial crisis until recently. Between 2013 and 2014 the dominant topic is indeed the one containing words such as “bond”, “purchase” and “debt” as well as “crisis” and “cut”. Incidentally, these topics are better captured by the SVM methodology than the SO, which would require the inclusion of *ad hoc* expressions to characterise monetary policy also in periods in which non-standard instruments (and vocabulary) are predominantly used.

Top 10 words for the most frequent topic identified by the LDA

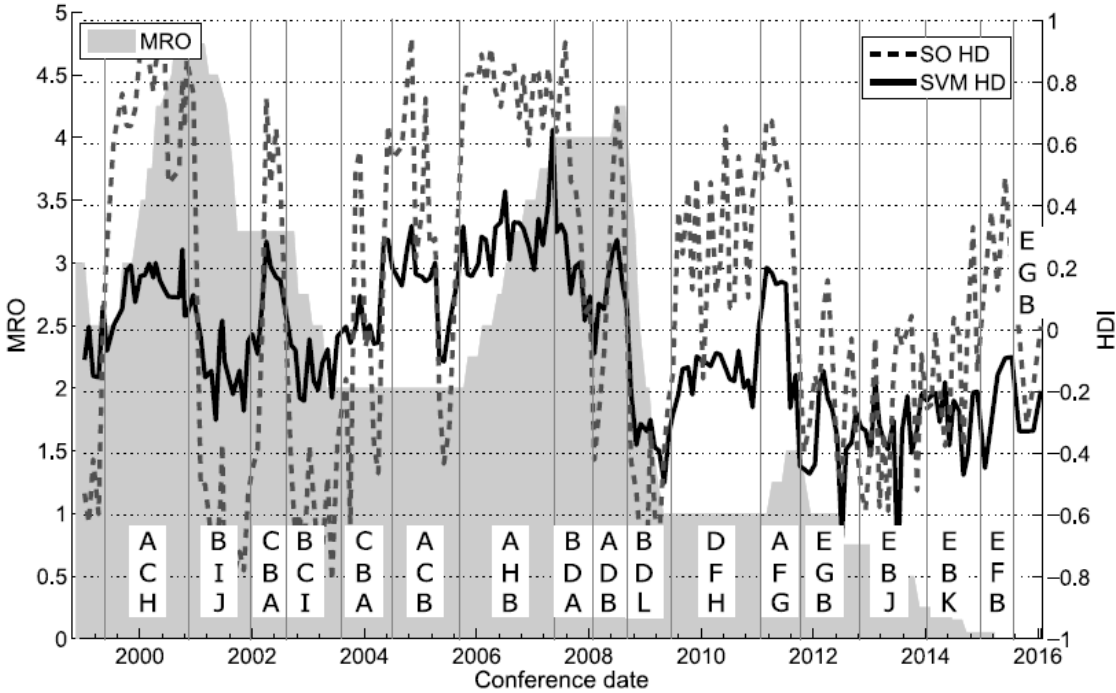
Table 2

Topics										
A	B	C	D	E	F	G	H	I	J	K
rate	rate	growth	Trichet	ECB	Greece	bond	price	cut	stock	rate
ECB	ECB	price	economy	Draghi	debt	market	growth	ECB	market	month
Trichet	cut	euro	market	bond	Greek	yield	remain	rate	share	Euribor
inflation	euro	economy	lend	crisis	ECB	debt	medium	point	index	market
price	bank	rate	record	cut	crisis	govern	risk	decision	bank	ECB
rise	zone	ECB	low	purchase	bond	Spain	inflation	move	expect	lend
bank	inflation	inflation	credit	monetary	default	purchase	stability	announce	rose	fixed
hike	expect	policy	recession	interest	bailout	crisis	expect	Thursday	trade	expect
interest	year	stability	financial	buy	plan	spread	develop	market	fell	overnight
increase	economy	recovery	crisis	debt	fund	Italy	continue	dollar	gain	EONIA

Note: Topics as identified by the LDA algorithm are indicated by the capital letter on top of each column

Figure 4. The HD index and most dominant topics in 13 periods

Figure 4



5. Conclusions

In this paper, we present the development of a numerical indicator that represents the perceived degree of hawkishness or dovishness based on news articles reporting on ECB monetary policy decisions. The evidence provided indicates that the HD index can bring value to central bank communication. On a methodological ground, data mining techniques such as SVM prove to be superior to develop an index measuring the tone of communication as they are more flexible and may offer more insightful information. Specifically, we show that it is possible to extract and analyse the most frequently used topics in the text data, that for the HD index confirms a shift in media focus from the standard interest rate setting to non-standard monetary policy instruments.

Apart from the results presented in this paper, the methodology can be easily applied to other communication outlets or extended to include other languages consistent with the multilingual nature of the Economic and Monetary Union to make it a fully-fledged tool to effectively monitor media reports and assessing effectiveness of ECB communication.

References

- Blei, D. M., A. Y. Ng and M. I. Jordan (2003), "Latent Dirichlet Allocation", *The Journal of Machine Learning Research*, 3, 993–1022.
- Fawcett, T. (2006), "An introduction to ROC analysis", *Pattern Recognition Letters*, 27, 861–874.
- Hansen, S. and M. McMahon (2016), "Shocking language: Understanding the macroeconomic effects of central bank communication", *Journal of International Economics*, 38th NBER International Seminar on Macroeconomics: S114–S133.
- Hayo, B., and M. Neuenkirch (2015), "Self-monitoring or reliance on media reporting: How do financial market participants process central bank news?", *Journal of Banking and Finance*, 59, 27–37.
- KOF Swiss Economic Institute (2007), "KOF Monetary Policy Communicator for the Euro Area". (<http://www.kof.ethz.ch>)
- Lucca, D.O. and F. Trebbi (2009), "Measuring central bank communication: an automated approach with application to FOMC statements". Technical Report National Bureau of Economic Research.
- Nechio, F. and R. Regan (2016), "Fed Communication: Words and Numbers", *FRBSF Economic Letters*, 2016–26.
- Provost, F. and T. Fawcett (2013), *Data Science for Business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media, Inc.
- Turney, P. D. (2002), "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 417–424.
- Weiss, S. M., N. Indurkha and T. Zhang (2010), *Fundamentals of predictive text mining*, Springer Science & Business Media.
- Wu, J. C., and F. D. Xia (2016), "Measuring the macroeconomic impact of monetary policy at the zero lower bound", *Journal of Money, Credit and Banking*, 48, 253–291.



Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC-Bank Indonesia Satellite Seminar on “*Big Data*” at the ISI Regional Statistics Conference 2017

Bali, Indonesia, 21 March 2017

Between hawks and doves: measuring Central Bank Communication¹

Stefano Nardelli,
European Central Bank

¹ This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.



EUROPEAN CENTRAL BANK

EUROSYSTEM

Stefano Nardelli

European Central Bank

Monetary Policy Strategy Division

Between Hawks and Doves Measuring Central Bank Communication

IFC – Bank Indonesia Satellite Seminar on “Big Data”

Bali, 21 March 2017

- Usually analysis is to evaluate **financial market developments** to check whether they are in line with what intended and avoid unintended consequences

«The policy stance may also be affected by a continued appreciation of the exchange rate»
(M. Draghi, 24 April 2014)

- or to extract what markets and analysts expect about **future monetary policy moves**

«While the governing council did not take any concrete action, the ECB president raised the rhetoric and said his central bankers would be “unanimous” in backing more radical measures, including quantitative easing, to cope with a “too prolonged a period of low inflation.” Analysts viewed the shift in tone as substantial ... » (Financial Times, 4 April 2014)

- However, the message delivered by the central bank **may not be perceived in the intended way** by stakeholders (financial markets, banks, public opinion)

«The markets got it wrong in forming their expectations. They did indeed have higher expectations than were there and that's why they reacted like they reacted but that was not our intention» (V. Constâncio, 3 December 2015)

«It is not the task of a central bank to correct the erroneous opinions of individuals ... the central bank has to do what it considers right » (E. Nowotny, 9 December 2015)

To assess to measure effectiveness of communication is crucial to develop **quantitative measures**

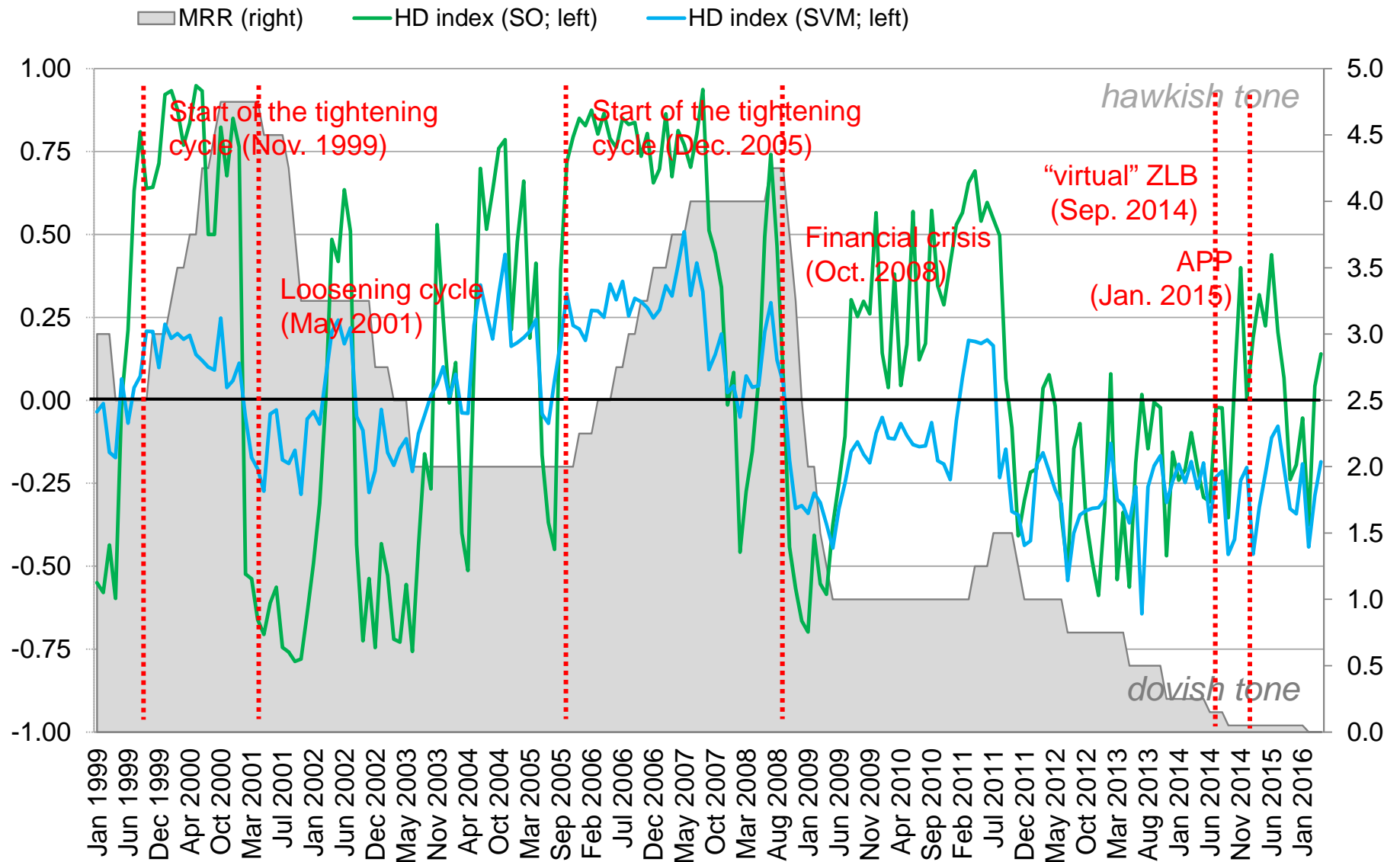
- *traditional approach*: “event study”, i.e. measure the impact on some market variables around some crucial communication events (e.g. press conference, speeches, press releases, etc.)
- *text analysis approach*: derive indicators directly on texts relevant for communication, issued either by CB or external watchers/stakeholders, using computational linguistic techniques

The HD index Two alternative methodologies

- The **HD index** is computed on articles and newswires reporting on ECB press conference
- First version computed using **Sematic Orientation** technique based on a fixed set or pre-determined word/expressions exogenously classified as either dovish or hawkish to determine the tone of a document
- Alternative version makes use of **Support Vector Machine (supervised machine learning algorithm)** to further reduce subjective interpretation
- This algorithm automatically looks for **patterns in text documents** to select the words with the highest discriminative power and determines the tone of a document based on them

- Main **data source** is **Factiva**, i.e. Dow Jones's global news database featuring nearly 33,000 sources (e.g. Dow Jones Newswires, The Wall Street Journal and Barron's). But *Financial Times* also included
- The **HD index** is computed on a subset of articles about the ECB press conference for a 3-day time window "around" the event (i.e. day before $t-1$, very day t and day after the pc $t+1$)
- The **time series** starts in January 1999 and covers the whole history of the euro (more than 9,000 articles in total)

The HD index



Question: **how to validate results?**

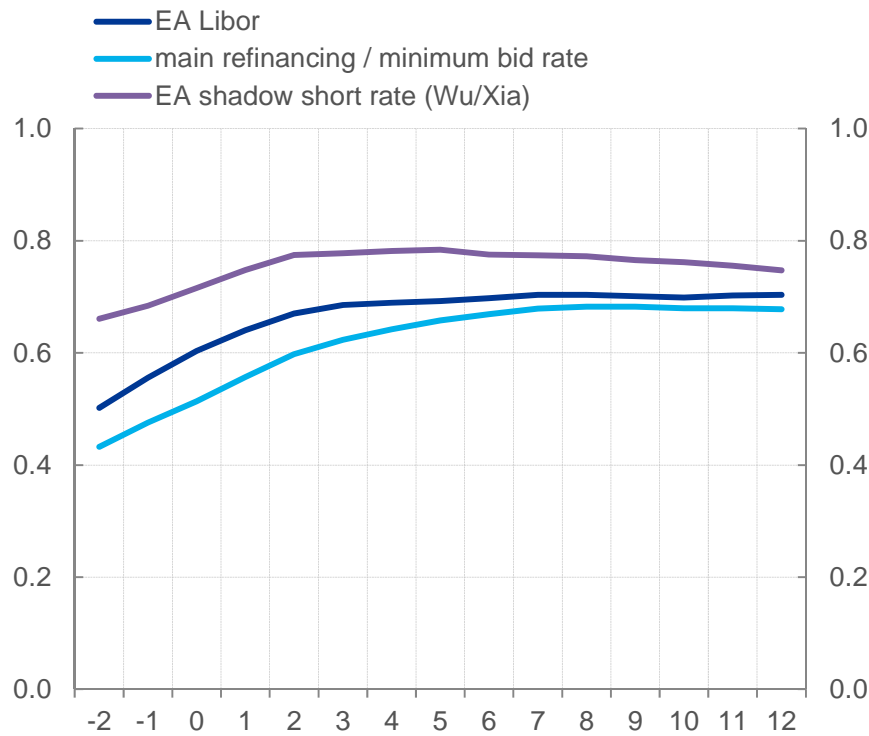
Three indicators are proposed:

- **Correlations** between HD and actual interest rates (“*are actual ECB monetary policy moves anticipated correctly by the media reports?*”)
- Identify topics on news dataset using a **topic classification model** (*unsupervised machine learning algorithm*) to assess whether media reports focus on most relevant topics
- To assess superiority of one method **performance metrics** (used increasingly in machine learning and data mining research)

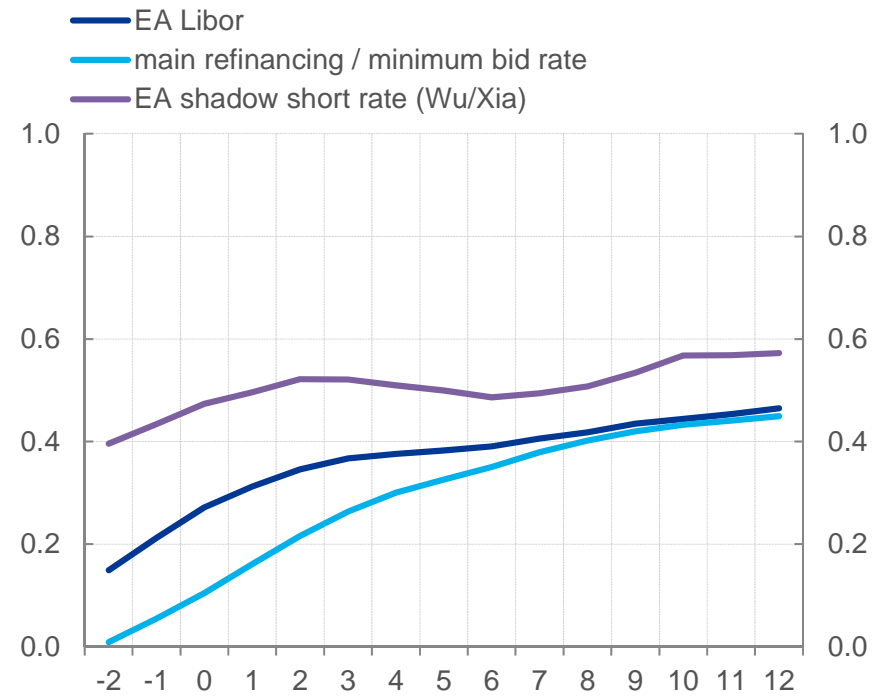
The HD index Leading properties (correlograms)

a. interest rate levels

SVM methodology

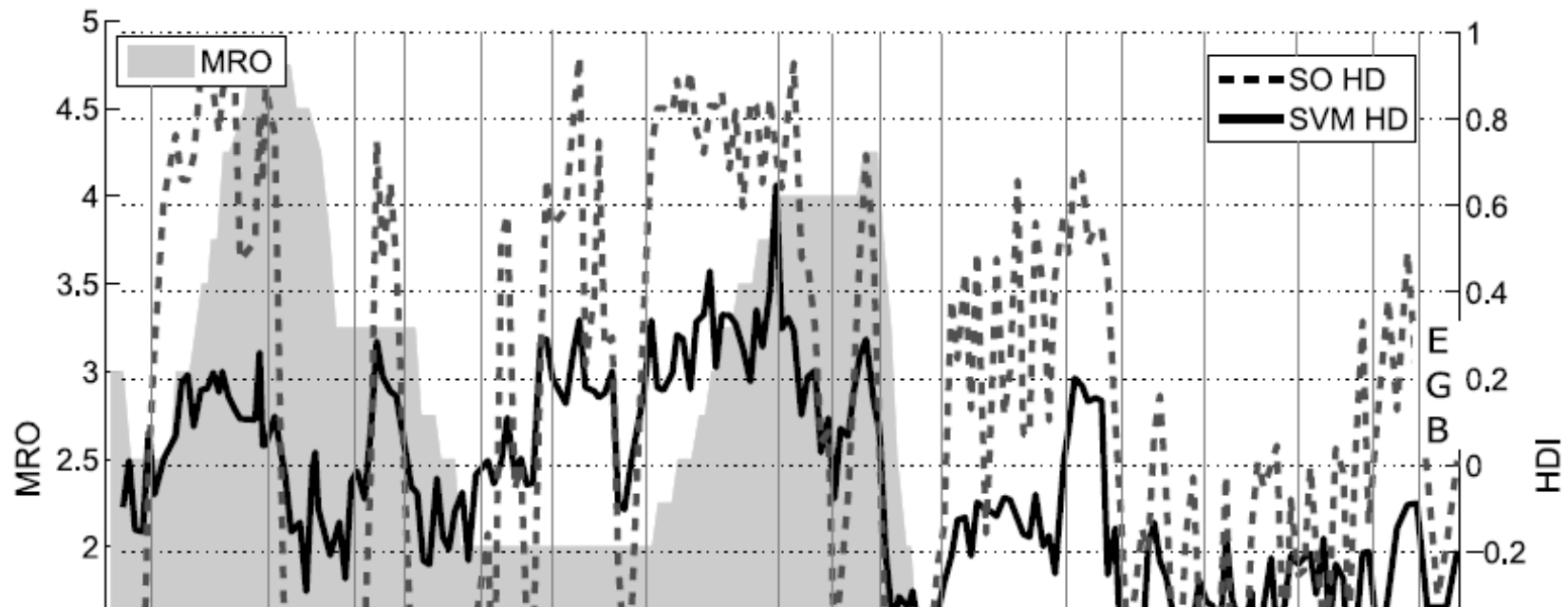


SO methodology



The HD index Topic analysis

The HD index and most dominant topics in 13 periods



A	B	C	D	E	F	G	H	I	J	K
rate ECB Trichet inflation price rise bank hike interest increase	rate ECB cut euro bank zone inflation expect year economy	growth price euro economy rate ECB inflation policy stability recovery	Trichet economy market lend record low credit recession financial crisis	ECB Draghi bond crisis cut purchase monetary interest buy debt	Greece debit Greek ECB crisis bond default bailout plan fund	bond market yield debt govern Spain purchase crisis spread Italy	price growth remain medium risk inflation stability expect develop continue	cut ECB rate point decision move announce Thursday market dollar	stock market share index bank expect rose trade fell gain	rate month Euribor market ECB lend fixed expect overnight EONIA

The HD index Performance analysis

- **Accuracy**, i.e. percentage of correctly classified observations
- **Area under the Receiver Operating Curve (AUC)**, i.e. measure to what extent positively labelled observations (hawkish) are ranked higher than negatively (dovish) labelled observations; AUC can deal with unbalanced distributions (Fawcett 2006)

Method	AUC	Accuracy
SO	69.24%	62.00%
SVM	98.55%	92.00%

- A quantitative approach to communication is possible
- Tools originally developed from computational linguistics may offer useful insights on CB communication and its perception by stakeholders (media)
- Specifically HD index is a sensible summary indicator about how communication around the ECB press conference is reproduced by media
- The SVM methodology proves superior to the “traditional” SO methodology