David-Jan Jansen and Jakob de Haan

An Assessment of the Consistency of ECB Communication using Wordscores

DNB Working Paper

De Nederlandsche Bank
Eurosysteem

An assessment of the Consistency of ECB Communication using Wordscores

David-Jan Jansen and Jakob de Haan *

An Assessment of the Consistency of ECB Communication using Wordscores[1]

David-Jan Jansen[a][2] and Jakob de Haan[a,b,c]

Version: 30 July 2010

[a] *De Nederlandsche Bank, Amsterdam, The Netherlands*

[b] *University of Groningen, The Netherlands*

[c] *CESifo, Munich, Germany*

Abstract

Wordscores uses word frequencies to extract information from texts with known policy positions. Wordscores uses this information to estimate the unknown policy positions of so-called virgin texts. We apply Wordscores to the ECB President's introductory statements following Governing Council meetings. We code policy positions of statements from the first three years of the Economic and Monetary Union (our reference texts) using various indicators of ECB communication as well as actual rate decisions. Treating introductory statements from 2002 to July 2009 as virgin texts, Wordscores is able to present a fairly accurate picture of ECB policy decisions during that period. The results also suggest changes in ECB communications occurred: using more introductory statements as reference texts improves the match between estimated positions and actual policy. Overall, we would characterize ECB communication during the first decade of EMU as internally consistent. At the same time, communication was flexible enough to adapt to changed circumstances.

---

1. Introduction

Over the past two decades, communication has become an important instrument for monetary policymakers. In line with this development, there has been a surge in empirical studies on central bank communication (Blinder, Ehrmann, Fratzscher, De Haan and Jansen 2008). Many of these studies refer to the communication policy of the European Central Bank (ECB). There is substantive evidence that ECB communications move financial markets in the intended direction (see, for instance, Ehrmann and Fratzscher 2007). Likewise, there is much evidence that ECB communications increase the predictability of interest decisions (see, for instance, Sturm and De Haan 2010).

An important requirement for communication to be effective is consistency. If central bankers too frequently change their vocabulary, it will be difficult for observers to properly infer the message that central bankers are trying to convey. This paper assesses this internal consistency of ECB communication. Put differently, we assess the similarities in the ECB's vocabulary over the years. How alike, for instance, was the way in which the ECB communicated its monetary policy stance in 1999 and 2009? This concept of consistency differs from the existing literature, which usually focuses on the match between words and subsequent policy decisions (Berger, De Haan and Sturm 2010). Another prominent take on consistency is the degree of dispersion in communications by various officials from the same central bank (Jansen and De Haan 2006, Ehrmann and Fratzscher 2007).

Although consistency is important, it is also good to note that a degree of flexibility may be necessary when communicating monetary policy. As circumstances change, communication is bound to look for new ways to express the motivations underlying policy decisions. From that perspective, too much emphasis on providing

consistency may mean too much rigour with respect to communication. Therefore, we also study potential changes in communication.

With respect to the ECB, there are a number of occasions on which changes in wording may have occurred. Some would argue that, as a new organisation, the ECB perhaps had to learn how to communicate effectively with the outside world. The first years of the Economic and Monetary Union (EMU) may, therefore, have constituted a learning period for the ECB. Second, after an internal assessment, the ECB clarified its monetary policy strategy in May 2003. Although this was not supposed to reflect a change in strategy, many observers interpreted the clarification as downplaying the importance of the role of money in the ECB's strategy (Berger et al., 2010). In any case, the terms 'economic analysis' and 'monetary analysis' were introduced as new elements in the introductory statements. Third, in order to provide additional relevant information in a timely matter, the ECB started to discuss the quarterly macroeconomic projections by ECB or Eurosystem staff in June 2004. Four times a year, therefore, the introductory statements will contain some information which they did not before mid-2004. Finally, the financial turmoil may have led communication to adapt since mid-2007. The turbulence in financial markets, followed by a severe economic downturn, may have led the ECB to find new words to describe the considerations underlying its monetary policy in this changed environment.

In this paper, we examine whether ECB communications have remained consistent over time, and whether possible changes in vocabulary have increased our understanding of monetary policy. We investigate these two issues by conducting the following thought experiment. Suppose someone had closely followed and analyzed ECB communications during the early years of EMU. How well would this person have

been able to understand monetary policy during subsequent years, *without heavily investing in any further kind of analysis*?

To conduct our thought experiment, we use the Wordscores methodology introduced by Laver, Benoît and Garry (2003). The idea underlying Wordscores is the following: from texts of which the policy position is known (the so-called reference texts), Wordscores extracts information using word frequencies, and this information is employed to estimate the policy positions of subsequent texts about which nothing is known (the so-called virgin texts). All that is used to infer the policy positions of virgin texts are the words contained in these texts, which are compared with the words observed in reference texts with a known policy position.

We apply the Wordscores methodology to the introductory statements by the President of the ECB at the press conference following the ECB's monetary policy meetings, as that is the ECB's most important communication device (Blinder et al., 2008). We start by using the introductory statements between 1999 and 2001 as reference texts and statements in subsequent years (2002 – July 2009) as virgin texts. Wordscores requires that the reference texts are coded. To do so, we use two coding schemes. First, we rely on various indicators of ECB communication developed by different authors. As an alternative, we use the ECB's actual policy decisions. To investigate changes in wording, we subsequently expand the time frame for the reference texts by one year. Thus, we are able to determine whether this enhances the ability of Wordscores to identify the policy positions of the virgin texts.

Given the reasons to expect that the ECB's communications may have changed over time, we were actually somewhat sceptic that the mechanical approach underlying Wordscores would be able to generate an accurate representation of ECB monetary policy. However, we find that solely on the basis of the introductory statements during

4

the first three years of the EMU, the Wordscores method is able to present a fairly accurate picture of the ECB's policy decisions in subsequent years. From that perspective, the ECB's communication as laid down in the introductory statements can be considered as quite consistent over the last decade.

At the same time, the evidence also suggests that ECB communications have changed: adding more years of introductory statements to the set of reference texts generally improves our understanding of monetary policy. From that perspective, it seems the introductory statements continued to provide new relevant information for central bank watchers. For instance, adding the statements from 2006 changes the relative occurrence of the word 'vigilance', enabling us to better match subsequent decisions to tighten monetary policy.

This paper proceeds as follows. Section 2 outlines the Wordscores methodology and discusses our particular application to ECB communication. Section 3 presents the results, and section 4 provides a concluding discussion.

2. Using Wordscores to analyze ECB communication

Laver et al. (2003) introduced the Wordscores methodology to estimate policy positions of political parties in the United Kingdom. In a nutshell, the idea underlying Wordscores is to estimate policy positions by comparing two sets of texts using the relative frequency of words. Prior knowledge on the first set of texts, called reference texts, is used to infer the policy positions of the texts in the second set, which are labelled virgin texts. All that is needed to infer the positions of the virgin texts are the words contained in these texts, which are compared with the words observed in reference texts with 'known' policy positions.

The Wordscores method has a particular take on analyzing texts. Laver et al. (2003, p. 312) note how the method treats 'texts not as discourses to be read, understood and interpreted for meaning – either by a human coder or a computer program applying a dictionary – but as collections of word data containing information about the position of the texts' authors on predefined policy dimensions.' The method we use in this paper, therefore, differs from papers that either devise their own coding schemes (Buliř, Čihák and Šmídková 2010) or use methods for automated content analysis (Bligh and Hess 2010). Laver et al. (2003) identify several benefits of the Wordscores methodology. First, it can be used for texts written in any language. Second, because it treats words unequivocally as data, Wordscores is also able to calculate confidence intervals around the estimated scores. Finally, there is no need for heavy human intervention, and Wordscores can be implemented efficiently using the programs that the authors have made available on the internet.[3]

To use Wordscores properly, the selection of appropriate reference texts is crucial. The reference texts are required to provide sufficient information on the policy dimension for which one would like to evaluate the virgin texts. Laver et al. (2003) identify three guidelines for the selection of reference texts. Most importantly, the reference texts should be used in the same context as the virgin texts. This implies that the reference texts should relate to a similar context as the virgin texts. It is precisely this key requirement that we use as the basis for studying consistency. In a sense, we are reversing the direction of the analysis. Wordscores requires similarity between reference texts and virgin texts in order to be effective. We will argue that, if it turns out

---

[3] Wordscores is implemented as a set of Stata programs. The programs can be downloaded at: http://www.wordscores.com. We present the technical details underlying Wordscores in the appendix. See also Lowe (2008) for a further discussion of the Wordscores method.

that reference texts are useful to interpret virgin texts, this leads to the conclusion that the two set of texts are internally consistent.

The other two guidelines are also quite intuitive. First, the reference texts should contain enough variation in policy positions, so that the complete spectrum of positions is covered. Finally, the reference texts should contain as many different words as possible. The more words the reference texts contain, the more likely it is that words from the virgin texts can be assigned scores.

We focus on the ECB's most important communication device, i.e., the President's introductory statement at the monthly press conference in which he reports on the decisions taken by the ECB's Governing Council. Following meetings of the Council, which typically take place on the first Thursday of each month, the ECB announces the monetary policy decisions at 13:45 (CET). Some 45 minutes later, at around 14:30, the ECB President and Vice-President hold a press conference that comprises two elements: a prepared introductory statement that contains the background considerations for the monetary policy decision, and a Questions & Answers (Q&A) part during which the President and the Vice-President are available to answer questions by the attending journalists. The introductory statement is understood to reflect the position and views of the Council, agreed upon on a word-by-word basis by its members (Ehrmann and Fratzscher 2009; Berger et al. 2010).

We use the ECB President's introductory statements from the early years of EMU in order to understand introductory statements in later years. To be more precise, as a first step we consider the period between January 1999 and December 2001 as the reference period, and use the introductory statements from that period to understand monetary policy between January 2002 and July 2009. As the ECB monetary policy has shown sufficient variation – in terms of the number of rate changes and their direction –

between 1999 and 2001, there is enough relevant information in the reference texts, which can be used to identify the policy positions of the virgin texts. Finally, we are able to use a respectable number of words as references. The introductory statements are of course much shorter than the party manifestos used by Laver et al. (2003), but our number of reference texts is much larger. Our baseline set of reference texts (1999 – 2001) contains well over 45,000 words. Therefore, the third requirement to apply Wordscores is also met.

Wordscores requires that the reference texts be coded. So, although Wordscores offers the ability to process a large body of communication quickly, there is still the need for initial content analysis. In that sense, the method differs somewhat from other methods that offer automated content analysis, such as Alceste (see, for instance, Bailey and Schonhardt-Bailey 2008) or Diction (see, for example, Bligh and Hess 2010). Because we wanted to focus on the specific use of Wordscores, we chose to use four indicators of ECB communication from the existing literature. These indicators are all based on the ECB President's introductory statements. We use the aggregate indicator of Berger et al. (2010), the updated indicator of Rosa and Verga (2007), the indicator of Ullrich (2009), and the KOF MPC Communicator index.[4] These various indicators code ECB communications on a numerical scale. Negative (positive) values are assigned to communications that are perceived as dovish (hawkish), and zero to those that appear to be neutral. Whereas some researchers restrict the coding to directional indications by using a scale between –1 and +1, others assign a finer grid that is at least suggestive of magnitude, e.g., by coding statements on a scale from -2 to +2.[5] As an alternative to these

---

[4] We downloaded the KOF-MPC data from
http://www.kof.ethz.ch/publications/indicators/communicator/en. URL last accessed on 25 March 2010. We thank Carlo Rosa, Giovanni Verga and Katrin Ullrich for making their data available.
[5] Using these indicators, Sturm and De Haan (2010) examine whether ECB communication adds information compared to the information provided by a Taylor rule model in which expected inflation and

four indicators, we employ the actual policy decisions by the ECB Governing Council. In doing so, we define the policy dimension ranging between easing (which we assign a score of -1) and tightening (which we assign a score of +1).[6]

The four indicators have been used previously for various purposes. Berger et al. (2010) examine the role of money in the policies of the ECB, using introductory statements of the ECB President at the monthly press conferences during 1999-2004. They construct an aggregate indicator as well as indicators of the economic and monetary analysis. We use their aggregate indicator as we want to include all possibly relevant statements. Rosa and Verga (2007) use the introductory statements to show that the predictive ability of these statements is similar to that implied by market-based measures of monetary policy expectations. To do so, they provide a glossary that translates the qualitative information of the press conferences into an ordered scale. Ullrich (2009) shows how the indicator measuring ECB rhetoric contributes to explaining inflation expectations formation. Finally, the KOF MPC offers a quantification of the statements of the ECB president made during the monthly press conference. The indicator is based on those parts of the introductory statements which reveal the Governing Council's assessment of developments which directly affect future price stability. As the indicators use different scales, we standardize each index, and then take the average of the standardized indicators. This average is the score for the reference texts. The policy dimension is defined as the number of standard deviations from a neutral policy stance.

We also use the Wordscores methodology to make inferences about changes in the ECB's vocabulary. As noted, Wordscores assumes that reference texts contain useful

---

output are used. They find that even though the indicators are sometimes quite different from one another, they add information that helps predicting the next policy decision of the ECB compared to the information provided by expected inflation and expected output growth.

[6] An interesting extension would be to use the *size* of the interest rate change as the policy dimension.

information for understanding virgin texts published at a later date. However, as Laver et al. (2003, p. 314) acknowledge with respect to the party manifestos that they analyze, this 'assumption is unlikely to be 100 percent correct, since the meaning and usage of words in party manifestos changes over time, even over the time period between two elections in one country.' As we have argued, the same may be true for ECB communications. To identify whether this is the case, we therefore gradually expand the window of introductory statements that we consider as reference texts. If ECB communication changes over time, the positioning of the virgin texts will then also change. As more information becomes available through the wording of the statements, we expect that the method will be better able to match actual interest rate policy. We check this by confronting the estimated policy positions of the virgin text with the actual developments in the ECB's monetary policy stance.

3. Results

First, we discuss how well the estimated policy positions obtained using Wordscores match with actual developments in ECB monetary policy. For clarity, we note that we are evaluating the match between the score for each introductory statement and the *contemporaneous* policy decision.[7] Figure 1 shows the results when we use the introductory statements between January 1999 and December 2001 as reference texts and the actual policy decisions to code them. The solid line shows the scores for the virgin texts, i.e., the introductory statements spanning the period January 2002 to July 2009. The figure also shows the 95% confidence interval around the scores. The bars indicate the timing and direction of changes in the ECB main refinancing rate. Figure 2

---

[7] Therefore, our results are not so much addressing the issue of predictability of policy decisions. We did evaluate how well the scores of the statements matched next month's decisions. In general, the match was not as good as in the contemporaneous case.

shows the results in case the reference texts are coded using the ECB communication indicators.

<Insert figures 1 and 2 about here>

At first glance, it is striking how the estimated scores for the introductory statements broadly follow the actual developments in ECB monetary policy. As such, the scores obtained from Wordscores are helpful in obtaining an indication of the ECB monetary policy stance over the years. Admittedly, the correlation is far from perfect, which indicates that the scores may not be the best forecasts. Nevertheless, using only information for the first three years of EMU goes a long way in understanding monetary policy in subsequent years. From this perspective, we conclude that ECB communication over the last decade may be considered as quite consistent.

Next, we address the possibility of changes over time. Tables 1 and 2 provide various metrics of how well the estimated policy positions match the contemporaneous policy decisions. The window for the reference texts is expanded by one year in subsequent steps. For instance, the last column shows results when we use the introductory statements between 1999 and 2006 as reference texts. Again, we relate the scores to the contemporaneous decisions.[8] The tables show the correlation between the scores of the virgin texts and the actual policy decisions (row 1). The remaining rows of the tables show indicators for the ability of the scores for the virgin texts to match policy decisions taken between January 2007 and July 2009. This sample includes the period when one of the changes in communication may have occurred. In Table 1 the reference texts are coded using actual policy decisions, while in Table 2 these texts are coded using the ECB communication indicators.

---

[8] Also for this moving window analysis, the results are similar to the contemporaneous case if we relate today's statements to next month's decisions. Results are available upon request.

<Insert tables 1 and 2 about here>

Although there is some variation, the overall impression is that adding additional introductory statements to the set of reference texts leads to a better correspondence between the scores and the actual development in the monetary policy stance. For instance, the correlation increases, although not monotonically. Also, the percentages of correct classifications, in most cases, increases over time, while the number of incorrect classifications decreases.

The last columns of tables 1 and 2 have interesting results. Adding the introductory statements from the year 2006 to the reference texts does not uniformly improve the correspondence between estimated positions and actual policy.[9] The correlations between the scores and actual interest rate policy decrease in comparison with ending the reference period in 2005. Also, the percentage of correct 'no change' decisions is lower. In contrast, decisions to tighten monetary policy are better matched. For the first time, we are able to correctly match two of the three decisions to tighten monetary policy in 2007 and 2008. Figures 3 and 4 further illustrate this point. The figures show the match between scores for the virgin texts and the actual policy decisions, using the introductory statements of the ECB President from 1999 to 2006 as reference texts. Compared to Figures 1 and 2, the estimated policy positions are more closely linked in case of decisions to tighten monetary policy, especially when the references texts are coded using the ECB communication indicators.

<Insert figures 3 and 4 about here>

---

[9] It would be interesting, therefore, to expand the set of reference texts further. Momentarily, this would leave us with too few interest rate changes to evaluate the estimated policy positions.

Table 3 presents another piece of evidence with respect to this finding. These results are based on a moving-window experiment, where we keep the reference period constant at three years. Once again, the first reference period is 1999 to 2001; the subsequent one is 2000 to 2002, and so forth. The last reference period begins in 2004 and ends in 2006. This period is the only one that includes 2006.[10] The results are based on using actual policy decisions to code the reference texts. The results from the earlier exercise are clearly supported. Once we include the introductory statements from 2006 to the reference texts, we find that the scores better match the actual policy decisions.

<Insert table 3 about here>

What could account for the specific results for 2006? In our view, the role of keywords such as 'vigilance' is one plausible explanation for the better match with respect to tightening decisions. In earlier work (Jansen and De Haan 2007), we discussed how this term (and variations thereof) was often used to signal the ECB's worry on rising euro area inflation expectations from early 2004 onwards. However, after December 2005, the use of 'vigilance' was increasingly seen as an indication of upcoming tightening of ECB monetary policy.[11] It seems plausible that adding 2006 to the set of reference texts changes the relative occurrence of vigilance, enabling us to better match tightening decisions in 2007 and 2008.

---

[10] For this exercise, one of the conditions for Wordscores to be effective is no longer automatically valid. Between the reference periods, there are now differences between the number of occasions on which interest rate change as well as the distribution between decisions to ease and tighten. For instance, the reference period between 2002 and 2004 only contains three occasions on which policy was changed. In each case, rates were eased.

[11] According to Bloomberg 'ECB president Jean-Claude Trichet has used the word "vigilant" to flag each of the six rate increases since late 2005' (Bloomberg News, 15 February 2007). Likewise, according to UBS: 'Trichet has made a practice of effectively pre-announcing hikes at the prior meeting with the use of the key "vigilant" phrase' (UBS FX Trade and Research, 9 January 2007).'

4. Conclusions

In this paper, we apply Wordscores - a method for textual analysis based on word frequencies - to the introductory statements by the ECB President. To use Wordscores, we code introductory statements from the early years of the Economic and Monetary Union using various indicators of ECB communication as well as actual policy decisions.

We see two contributions from our analysis. First, we introduce Wordscores to the literature on central bank communication. From a pure methodological point, Wordscores may be an interesting addition to the researcher's toolkit. The underlying algorithm incorporated in Wordscores may help researchers in more efficiently analysing vast amounts of communications by central banks. At the same time, Wordscores still requires some a priori coding for the reference texts, which could still be time-intensive. Therefore, Wordscores is more likely a complement to, rather than a substitute for other methods of automated content analysis.

We also introduce Wordscores, as we think it can be used to analyze consistency. Therefore, this paper's second contribution is an assessment of the consistency of ECB communication during the first ten years of its operations. In contrast to much of the existing literature, we focus on the internal consistency of ECB communication over the years. There are two main conclusions. First, by combining a thorough analysis of ECB communication from the early years of the EMU with Wordscores, we are able to present a fairly accurate picture of ECB monetary policy in subsequent years. From that perspective, the ECB's communication as laid down in the introductory statements can be seen as quite consistent over the last decade. This point becomes especially clear if we again consider the simplicity of the Wordscores methodology. We only use information on the number of times particular words appear in ECB statements. Using

more sophisticated methods would probably only increase the matching between estimated and actual policy decisions.

Although consistency is important in communication, there is, in all likelihood, also a need for flexibility. Therefore, we also try to assess the role of possible changes in communication. Adding more introductory statements to the set of reference texts helps, in general, to improve our understanding of monetary policy in subsequent years. From that perspective, the introductory statements seem to have provided new relevant information for central bank watchers. Overall, we conclude that ECB communication during the first decade of EMU was characterized by consistency as well as flexibility.

References

Bailey, A. and C. Schonhardt-Bailey. 2008. 'Does Deliberation Matter in FOMC Monetary Policymaking? The Volcker Revolution.' *Political Analysis* 16 (4): 404-427.

Berger, H., J. De Haan, and J.-E. Sturm. 2010. 'Does Money Matter in the ECB Strategy? New Evidence Based on ECB Communication.' *International Journal of Finance and Economics.* Forthcoming.

Bligh, M. and G. D. Hess. 2010. 'Deconstructing Alan: A Quantitative Assessment of the Qualitative Aspects of Chairman Greenspan's Communication' In *Central Bank Communication, Decision-Making and Governance,* ed. P. L. Siklos and J.-E. Sturm. Cambridge, Mass.: MIT Press. Forthcoming.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. De Haan and D. Jansen. 2008. 'Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence.' *Journal of Economic Literature* 46(4): 908-943.

Buliř, A., M. Čihák and K. Šmídková. 2010. 'Writing Clearly: ECB's Monetary Policy Communication.' In *Central Bank Communication, Decision-Making and Governance,* ed. P. L. Siklos and J.-E. Sturm. Cambridge, Mass.: MIT Press. Forthcoming.

Ehrmann, M. and M. Fratzscher. 2007. 'Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness?' *Journal of Money, Credit and Banking* 39(2-3): 509-541.

Ehrmann, M. and M. Fratzscher. 2009. 'Explaining Monetary Policy Decisions in Press Conferences.' *International Journal of Central Banking* 5(2): 41-84.

Jansen, D. and J. De Haan. 2006. 'Look Who's Talking: ECB communication during the first Years of EMU.' *International Journal of Finance and Economics* 11(3): 219-228.

Jansen, D., and J. De Haan. 2007. 'The Importance of being vigilant: has ECB communication influenced euro area inflation expectations.' De Nederlandsche Bank Working Paper No. 148.

Laver, M., K. Benoît, and J. Garry. 2003. 'Extracting policy positions from political texts using words as data.' *American Political Science Review* 97(2): 311-31.

Lowe, W. 2008. 'Understanding Wordscores.' *Political Analysis* 16(4): 356-371.

Rosa, C., and G. Verga. 2007. 'On the Consistency and Effectiveness of Central Bank Communication: Evidence from the ECB.' *European Journal of Political Economy* 23(1): 146-75.

Sturm, J.-E. and J. De Haan. 2010. 'Does central bank communication really lead to better forecasts of policy decisions? New evidence based on a Taylor rule model for the ECB.' *Weltwirtschaftliches Archiv/Review of World Economy*. Forthcoming.

Ullrich, K. 2008. 'Inflation expectations of experts and ECB communication.' *North American Journal of Finance and Economics* 19(1): 93-108.

Table 1: Accuracy when using actual policy decisions to code statements

|	Reference periods ends in:					
	2001	**2002**	**2003**	**2004**	**2005**	**2006**
Correlation	0.26	0.33	0.21	0.43	0.65	0.52
% correct ease	0.0	0.0	16.7	50.0	50.0	33.3
% correct tight	0.0	33.3	0.0	0.0	0.0	66.7
% correct no change	75.0	80.0	75.0	80.0	85.0	70.0
% incorrect ease	NA	NA	0.0	40.0	50.0	50.0
% incorrect tight	100.0	80.0	100.0	100.0	NA	66.7
% incorrect no change	37.5	33.3	34.8	27.3	26.1	26.3

Note: The correlation is measured between an indicator (measured on a scale of -1, 0, 1) of actual decisions and the Wordscores for the introductory statements. The correlation is computed over the entire sample of virgin texts. All other measures are based on an evaluation for the period January 2007 to July 2009. The percentage of correct classifications is conditional on the actual decisions, while the percentage of incorrect classifications is conditional on classifications. The top row shows the last year of the reference period. The classification is determined by checking the scores for the texts. When the score is higher (lower) than +0.5 (-0.5), the predicted value is equal to +1 (-1), whereas in between the predicted value is 0.

Table 2: Accuracy when using the average of four ECB communication indicators to code statements

	Reference periods ends in:					
	2001	**2002**	**2003**	**2004**	**2005**	**2006**
Correlation	0.38	0.34	0.28	0.56	0.65	0.62
% correct ease	33.3	16.7	16.7	50.0	50.0	33.3
% correct tight	0.0	0.0	0.0	0.0	0.0	66.7
% correct no change	90.0	90.0	85.0	95.0	95.0	75.0
% incorrect ease	33.3	0.0	0.0	25.0	25.0	0.0
% incorrect tight	100.0	100.0	100.0	NA	NA	71.4
% incorrect no change	28.0	30.8	32.0	24.0	24.0	25.0

Note: The correlation is measured between an indicator (measured on a scale of -1, 0, 1) of actual decisions and the Wordscores for the introductory statements. The correlation is computed over the entire sample of virgin texts. All other measures are based on an evaluation for the period January 2007 to July 2009. The percentage of correct classifications is conditional on the actual decisions, while the percentage of incorrect classifications is conditional on classifications. The top row shows the last year of the reference period. The classification is determined by checking the scores for the texts. When the score is higher (lower) than +1.0 (-1.0), the predicted value is equal to +1 (-1), whereas in between the predicted value is 0.

Table 3: Accuracy when using moving windows of three years

	Reference periods:					
	99-01	**00-02**	**01-03**	**02-04**	**03-05**	**04-06**
Correlation	0.50	0.59	0.64	0.42	0.61	0.44
% correct ease	50.0	33.3	83.3	50.0	33.3	0.0
% correct tight	0.0	0.0	0.0	0.0	0.0	66.7
% correct no change	80.0	90.0	100.0	100.0	95.0	75.0
% incorrect ease	25.0	33.3	0.0	0.0	0.0	100.0
% incorrect tight	100.0	100.0	NA	NA	100.0	66.7
% incorrect no change	27.3	28.0	16.7	23.1	26.9	31.8

Note: The correlation is measured between an indicator (measured on a scale of -1, 0, 1) of actual decisions and the Wordscores for the introductory statements. All measures are based on an evaluation for the period January 2007 to July 2009. The percentage of correct classifications is conditional on the actual decisions, while the percentage of incorrect classifications is conditional on classifications. The top row shows the last year of the reference period. The classification is determined by checking the scores for the texts. When the score is higher (lower) than +1.0 (-1.0), the predicted value is equal to +1 (-1), whereas in between the predicted value is 0.

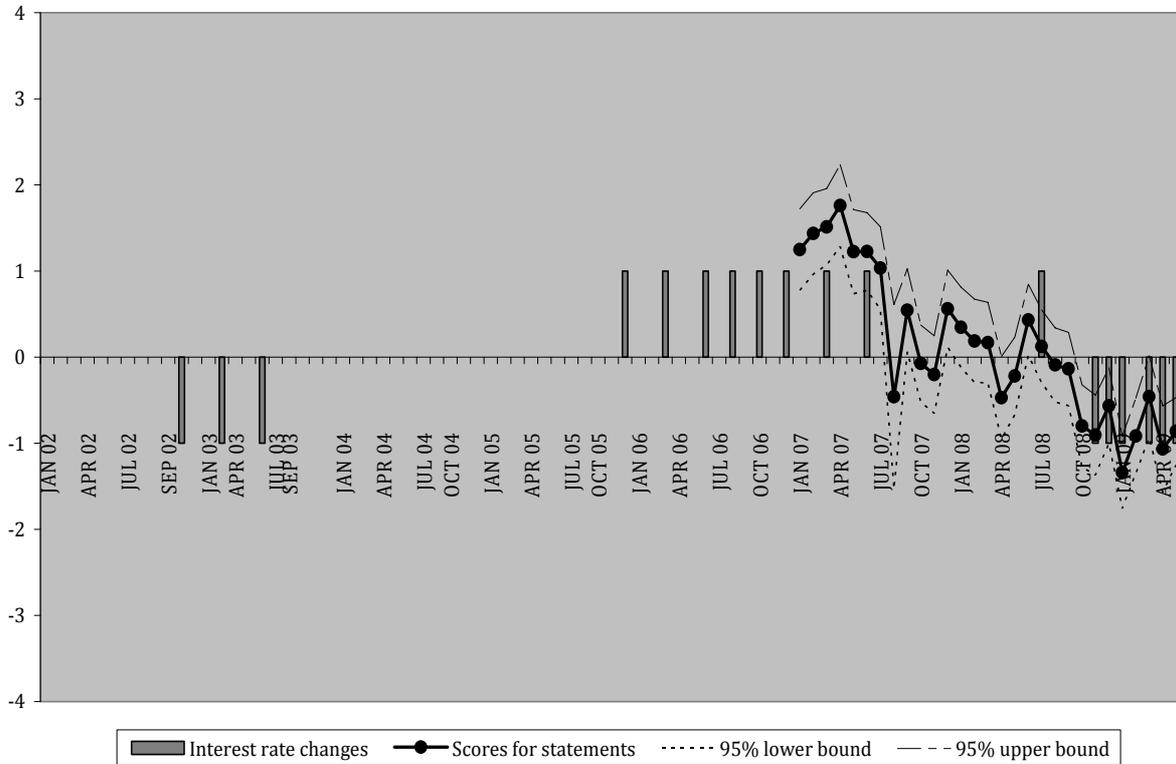Figure 1. Results for 2002-2009 using policy decisions as anchors
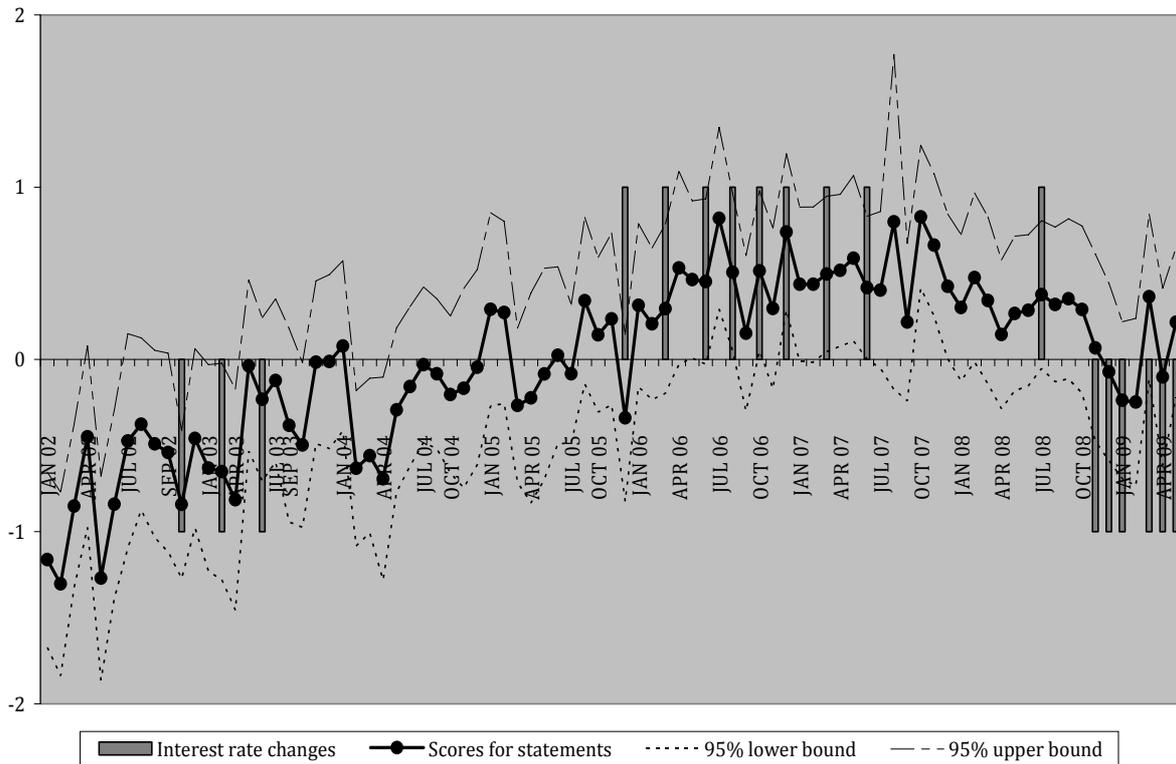
Note: The dots denote the scores for the introductory statements between January 2002 and July 2009. The dotted lines denote the 95% confidence interval around these estimates. The bars denote the change in the ECB policy rate, where positive values indicate tightening and negative values denote easing. If there are no bars for a certain date, policy rates remained unchanged. The reference texts are taken for the period 1999-2001 and coded using actual ECB policy decisions.

Figure 2. Results for 2002-2009 using indicators as anchors
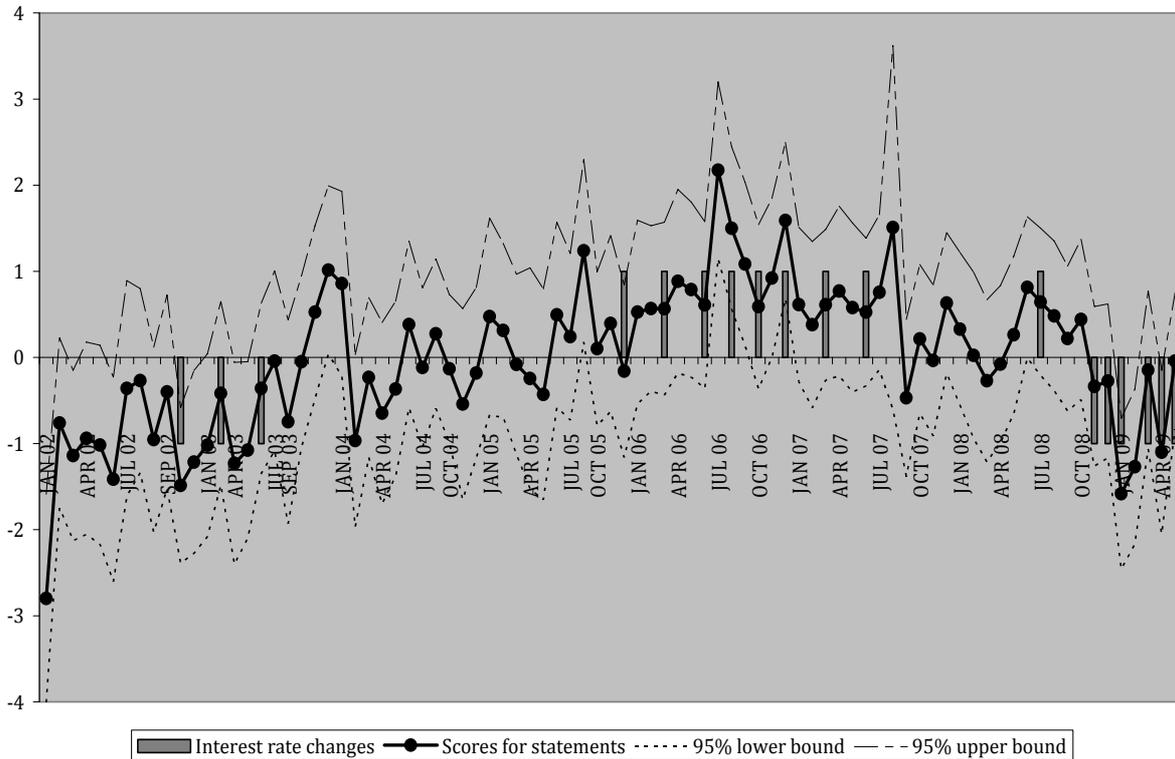
Note: The dots denote the scores for the introductory statements between January 2002 and July 2009. The dotted lines denote the 95% confidence interval around these estimates. The bars denote the change in the ECB policy rate, where positive values indicate tightening and negative values denote easing. If there are no bars for a certain date, policy rates remained unchanged. The reference texts are taken for the period 1999-2001 and coded using four ECB communication indicators from the literature.

Figure 3. Results for 2007-2009 using policy decisions as anchors
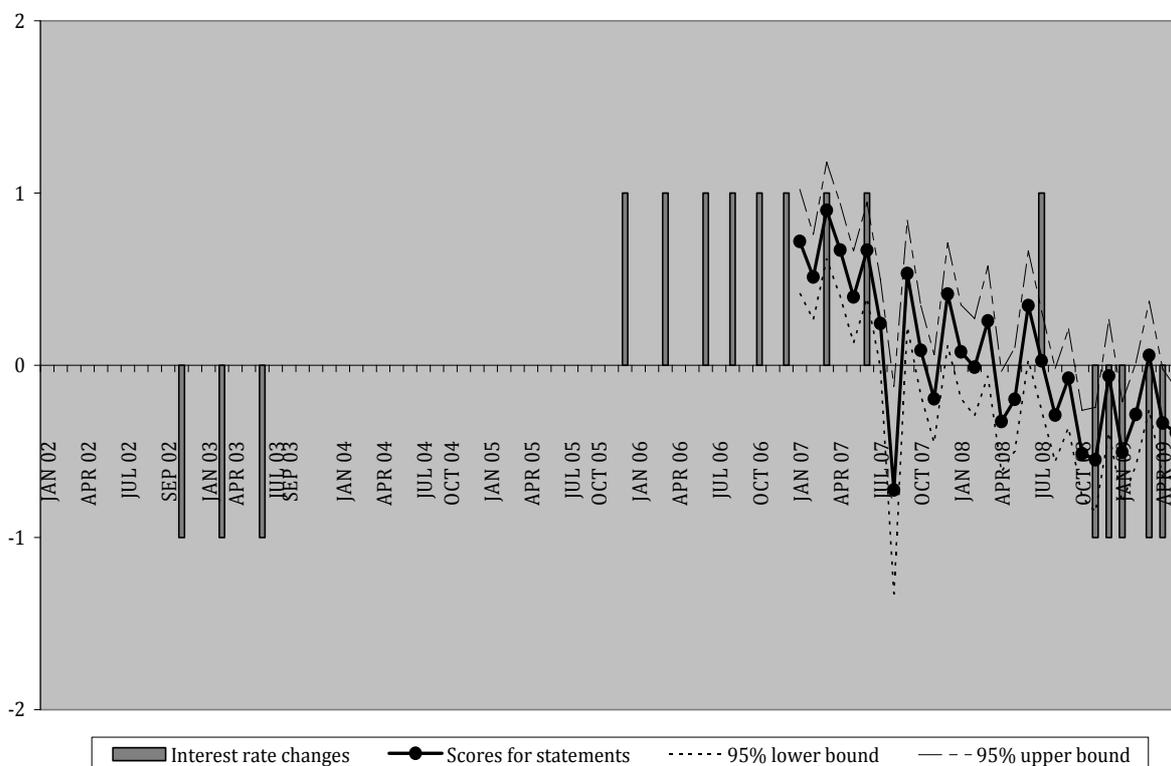
Note: The dots denote the scores for the introductory statements between January 2007 and July 2009. The dotted lines denote the 95% confidence interval around these estimates. The bars denote the change in the ECB policy rate, where positive values indicate tightening and negative values denote easing. If there are no bars for a certain date, policy rates remained unchanged. The reference texts are taken for the period 1999-2006 and coded using actual ECB policy decisions.

Figure 4. Results for 2007-2009 using indicators as anchors

Note: *The dots denote the scores for the introductory statements between January 2007 and July 2009. The dotted lines denote the 95% confidence interval around these estimates. The bars denote the change in the ECB policy rate, where positive values indicate tightening and negative values denote easing. If there are no bars for a certain date, policy rates remained unchanged. The reference texts are taken for the period 1999-2006 and coded using four ECB communication indicators from the literature.*

Wordscores proceeds in three steps[12]:

Treatment of reference texts:

1. There is a set R of reference texts, each of which has a position on the policy dimension d.

2. The position on this dimension of individual texts – denoted by r – is denoted by A_{rd}.

3. The relative frequency, as a proportion of the total number of words in a text, of each unique word w in text r is denoted by F_{wr}.

4. Wordscores uses the matrix of F for all unique words in a text to estimate the conditional probability that one is reading text r given that one is reading word w.

5. This probability is calculated as:

   $P_{wr} = F_{wr} / \sum_r F_{wr}$

6. The matrix of the conditional probabilities for all unique words is then used to produce a score S_{wd} for each word on dimension d. This score is the expected position on dimension d, given that one is reading w.

7. This score is the average of *a priori* scores for the reference texts weighted by the probabilities P. It is calculated as:

   $S_{wd} = \sum_r (P_{wr} * A_{rd})$.

---

[12] This appendix is based on Laver et al. (2003), pp. 315-316.

Treatment of virgin texts

1. First, the relative frequency of each word in the virgin texts is calculated. This frequency is denoted by F_{wv}.

2. The score of any virgin text v on dimension d is the mean dimension score of the scored words weighted by the frequency of the scored words. It is computed as:

$$S_{vd} = \sum_w (F_{wv} * S_{wd})$$

Transformation of raw virgin scores

1. A final technical issue is the fact that common words in the reference texts tend to be assigned the mean overall scores of the reference texts.

2. Therefore, raw virgin texts scores tend to be more clustered than the reference texts.

3. To deal with this issue, Laver et al. (2003) propose a transformation to guarantee that the scores of the virgin texts have the same dispersion metric as the reference texts. The results presented throughout this paper are the transformed scores.

4. The transformation is done by computing:

$$S_{vd}{}^* = (S_{vd} - S_{\acute{v}d})(SD_{rd}/SD_{vd}) + S_{\acute{v}d}$$

where $S_{\acute{v}d}$ is the average score of virgin texts, and SD_{rd} and SD_{vd} are the sample standard deviations for reference and virgin text scores.

Previous DNB Working Papers in 2010