Lund University
School of Economics and Management

LUND
UNIVERSITY

Evaluating the Riksbank's Transparency using methods of Textual Analysis

Bachelor Thesis, autumn term 2016
Author: Christian Barsoum
Supervisor: Fredrik N.G. Andersson

Abstract

In this paper we develop a method for measuring transparency as a factor of document similarities between the executive and operational branches of the Swedish Riksbank. We do this by usage of NLTK within the Python-framework. We also propose a ray of different methods for textual analysis, such that they are able to capture the elements of communication that we are interested in.

Keywords: Central bank communication, text analysis, transparency.

Table of contents

1. Introduction

This essay explores the efficiency by which the Swedish central bank, Riksbanken, communicates its intentions and observations concerning monetary policy, both internally and externally. We argue for the importance of transparency in the Riksbank's work, and highlight the major role the institution's communication plays in societal functions, such as market stability. The inherent value of this lies in, amongst others, the public's ability to scrutinize and analyse the Riksbank's functionality, to monitor its objectives, and the performance regarding their execution.

We carry out our study using various methods of textual analysis, by which we attempt to quantify communication. We do it on documents published by the Riksbank. They are the *Minutes of the Executive Board's monetary policy meetings* and the *Monetary Policy Reports*. This is achieved by the usage of a natural language processing package available to the Python-programming language. Python is an object-oriented programming language, meaning it is able to encase and alter different forms of data into 'objects', making it easy to handle and analyse large sets of data (Kindler and Krivy, 2011). Accompanying Python will be a package called the Natural Language Toolkit (Bird, et al., 2015). NLTK provides us with the functionality necessary to perform textual analysis on greater quantities of data. The mere counting of terms will fair us no better than attempting to read all the corpus ourselves, and try to differentiate between the documents by hand. Instead, we can use far more sophisticated methods of textual analysis allowing us to, for example, weigh certain words and visualise the difference between documents as a whole – methods which are provided to us by the NLTK. We have chosen to approach the issue using this method because of the practical simplicity of its implementation into both the Python framework, and the relevant documents.

Our approach stems from two premises. The first is that each communicated piece of information is expressively reflective of a level of transparency. We assert that the information contained in all monetary policy reports are extracted from the minutes of the executive board. Their interpretation of said minutes provide a measure of understanding, regarding current and expected states of the economy. Thus, transparency is measured as the rate of correspondence between the two types of documents. Analytical emphasis is placed on topic selection and the magnitude of attention put on each individual topic.

The second premise is that the level of external transparency equals that of the internal. All public information of the Riksbank's intentions are retrieved from the relevant published

sources. As the public partakes in the monetary policy reports, it is inadvertently confronted with the level of internal transparency of the Riksbank. As per our first premise, the level of information received by the public corresponds to that which is transferred from the minutes into the reports. Thus, the external level of transparency, from the Riksbank to the public, matches the internal level of transparency, from the minutes to the reports.

We motivate this approach due to its unambiguously defined parameters – namely the corpus, our two types of documents. The ease with which we define the data to collect, and still come to justified conclusions, differs from previous attempts at measuring central bank transparency. The following is an array of such endeavours, as laid out by the European Central Bank's survey on communication and monetary policy (Blinder, et al. 2008). In many instances it has been the case that studies, attempting to measure central bank transparency, have landed on a measure of impact of the communication on various factors, rather than an assessment of the communicated information instead. One such measure is the effect on *market volatility* caused by communication from the central bank, in which financial variables, such as asset returns, are determined to have moved in certain ways as a response to monetary policy communication (Kohn and Sack, 2004; Connolly and Kohler, 2004; Reeves and Sawicki, 2007). Another approach has been to *quantify* communication in order to measure its effects on financial variables, such as exchange rates (Jansen and De Haan, 2005), and economic outlook (Ehrmann and Fratzscher, 2007). Although not primarily concerned with communication, similar attempts at quantifying the qualitative aspects of central banks - here transparency and independence - has yielded results which support the idea that a trait such as communication indeed can, and ought to, be measured (Dincer and Eichengreen, 2013).

We choose to approach the issue of central bank transparency in our defined manner due to its comparably approachable nature. As we have determined monetary policy documents to be the only necessary factors for measuring transparency, the process of gathering data becomes a simple practical matter of executing the correct methods of analysis. Unlike previous literature, we do not concern ourselves with the impact of communication on external variables, but instead only on communication itself. Thus, our only source for data are the two types of documents in mention.

The essay will constitute the following parts: first (2) is an overview of different models for transparency, and the history of communication within central banking operations. This follows together with an outline of how data processing and information retrieval has developed over time as sophisticated tools of qualitative analysis. In the following section (3)

the methodology used is explained, such as the technicalities of interpreting the data results. Thereafter (4), we present the results in appropriate tables and graphs. Later (5), we discuss the results and provide an insight of their implications, as well as suggest possible explanations for the produced behaviours. Finally (6), the essay is summarized and its central points made clear.

2. Central Banking Communication

There are countless of ways of approaching the issue of transparency, and the opinion on what to look for has changed over time. Over the years, communication has come to be regarded as an essential tool in the monetary policy pursuits of the central banks. This stems from the gradual development of support towards a transparent public sector, and the need for an open-discussion norm in politics. Thus, monetary policy planning has shown a gradual progression towards a more open setting in their conducts, in favour of predictability over obscurity. Over the years, this has been backed up with major academic support.

2.1 Models of transparency

Presently, the dominant disposition of the central banks is to regard communication and transparency as vital components of their practices and responsibilities (Ehrmann and Fratzscher, 2005; Sveriges Riksbank, 2016a). They recognise the influence and control their words hold in society, and deem it as equally important as any effective policy plan.

Though that is hardly disputable, the methods by which the central banks carry out their communication has proven to be of vital concern. For instance, one might argue that it is important to reflect the diversity of opinions of the executive board members, as it provides a broader understanding of the current and anticipated economic conditions (Bernanke, 2004). Whilst others argue for the need of a clear communications strategy, with a few select ideas being developed, as to not cause confusion amidst public and private actors in the market (Issing, 2005).

At present time, there appears to be little dispute over the influence that central banking communication has on market expectations. In fact, monetary policy is said to effectively be a "management of expectations" (Woodford, 2004; 2005). The case for central bank transparency is also being proposed as having major beneficial effects on policy design overall, increasing the possibility of central banks managing monetary policy at the optimal societal level (Faust and Svensson, 2001). The major driving force behind this is the fact that

transparency creates opportunity for predictability, allowing economic agents to adapt accordingly to new interest rates (Apel and Vredin, 2007). Thus, we can conclude that transparency has a major role in current central banking affairs. The idea of establishing a norm for transparent communication opens opportunities for not only optimizing monetary policy decisions, but also invigorating the public's trust in forecasts.

2.2 Chronology of central banking communication

The role of the central bank lies in administering the money supply, and thus altering several economic factors such as interest rate, inflation and more. However, traditionally, their affairs and operations have not always been subject to a transparent way of working, making their agendas practically unknown or hard to understand for most of their existence. As quoted in a paper by the European Central Bank on monetary policy communication, prior to the recent '90s:

> *Central Banking […] thrives on a pervasive impression that [it] […] is an esoteric art. Access to this art and its proper execution is confined to the initiated elite. The esoteric nature of the art is moreover revealed by an inherent impossibility to articulate its insights in explicit and intelligible words and sentences.* (Blinder, et al. 2008)

The general consensus of the 1970s and '80s were that central banks, in contrast to todays' views, saw secrecy as an advantage to their planning and policy decision-making (Kang, et al., 2013). It was thought that keeping the public out of the operations of the central banks also prevented it from becoming a branch of government, ruled by political interest (Mishkin, 2004). This ensured the central bank's autonomy and made it possible to plan for long-term monetary policy decisions. Per the report of Kang et al. (2013), we learn that it was not until the year 1990 that the first central bank, that of New Zeeland, adopted an explicit inflation targeting. This act was not followed up on until a decade later by Norway, Sweden, and the United States in 2001, 2007 and 2012 respectively.

Unsurprisingly, we may within reason assume the general adoption of public disclosure has been a slow, and tedious process. As mentioned, it has always been in the interest of the central banks to remain self-governing, for the sole purpose of dividing the responsibility of monetary and fiscal policy. The emergence of conflicting interests between fiscal goals, and monetary ones, have always been reason for uncertainty and division in policy construction. This, added

with the problem of short-sighted planning on behalf of ruling political parties, have been reasons for desiring autonomy in the sphere of monetary policy.

Though perhaps contradictory at first sight, a key element to democratic functionality in a society ought to encourage a sacrosanct central bank. This can be seen as a safe-check mechanism of society, ensuring that extreme political interests may not interfere with the way monetary policy is being conducted. Perhaps it was this force of reasoning that, though initially prevented central banks from disclosing information about their activities, led them to gradually become more open. As they slowly realized the benefits of disclosing their intentions and predictions to the public, they simultaneously ensured their impact on the state of the economy. As the market has become dependent on the central bank's anchoring, prediction, and neutrality towards fiscal policies, they ensure a dependability of the public for them to remain unassociated with government.

Thus, from a practical standpoint, we can witness the gradual increase of support for greater levels of transparency in the operations of the central banks. In 1991, Marvin Goodfriend proposed an advancement in the utilization of interest rate targeting, as conducted by the Federal Reserve, for the purpose of promoting a more practicable agenda from the central banks (Goodfriend, 1991). Prior to this, communicating intentions were not of primary concerns to the central banks. He suggested the strategic advantages of a more accessible monetary policy-framework, in that it allows for higher predictability in the market spheres. This would in turn open grounds to a more cohesively operated economy.

The newly introduced inflation targeting had indeed increased the transparency of the central banks, and with it followed the effects as proposed by Goodfriend, and others (Mishkin, 2014). In addition to this, Woodford (2012) speculated that by anchoring inflation central banks allowed for the market to adjust to the developing rates of inflation, increasing productivity and market efficiency as a whole.

A decade later than Goodfriend, Woodford settles the importance of the matter when he affirms, in accordance with Goodfriend, that "*successful monetary policy* [is about affecting] *the evolution of market expectations*" (Woodford, 2001).

This is complemented by William Poole (2001) in the same year, claiming "*the presumption must be that market participants make more efficient decisions […] when markets can correctly predict central bank actions*". Transparency has thus become a key element of monetary policy in the new millennia, setting the agenda of how to reform numerous central banks across the world. According to Blinder et al. (2008), the driving force for transparency might have been

due to an increased sense of accountability for independent institutions. This ties together with the previous notion of independence being the product of transparency, and transparency being the compensation for its independence.

Subsequently, following the turn of the century, we are witnessing a revision in the narrative of how central banks ought to be managed, with emphasis on public disclosure and transparency being encouraged. Attention has gradually been shifted towards communication as a tool itself in the fulfilment of policy rulings. This trend follows the increased interest, and influence, of behavioural economics in recent decades as a means of complementing our understanding of economic dynamics. Market forces are persistently dictated by public perception and reaction to economic stimuli, such as new monetary targeting strategies (Bernanke & Mishkin, 1992).

As a natural response to this unfoldment in history, i.e. the increased importance of communicating intentions, the Swedish Riksbank recently declared a new policy titled: *The Riksbank's communication policy* (Sveriges Riksbank, 2016a). In it lies the support for the historical direction central banking operations have been moving towards. They clearly state the importance of using communication as a strategic tool for the purposes of the Riksbank, the reasons being:

- The possibility of public and private sectors to examine and evaluate the work of the Riksbank.

- Maintaining a high level of confidence for the operations of the Riksbank.

- Contribute to the understanding of monetary policy, and effectively adjusting predictions about it.

- Improve the internal level of transparency within the Riksbank, ensuring all employees are on the same track and effectively coordinating their work (Sveriges Riksbank, 2016a).

It is clear that the attitude of central banks, here the Riksbank, have most certainly changed during the last three decades. Time has shown the importance of a transparent practice in establishing grounds for efficient monetary policy creation. This does not only concern the public understanding and predictions but also, as the Riksbank points out, the internal

coordination is at least as important. Moving on from here, we ought to expect the central banks of the world to continue to improve their means of communication and cultivate a more transparent behaviour. The Riksbank's communication policy ought to only be one in a number of future progressions towards meeting more developed means of communication.

3. Methodology

There is no given definition of transparency. One possible definition is presented here, and concerns the level of correspondence between documents from the executive branch of the Riksbank, and the monetary policy divisions. We explore the efficiency by which the internal (executive) discussion is reflected in the publicly aimed documents, the monetary policy reports.

In order to evaluate the transparency of the Riksbank, we need to define our approach and motivate its use. For our purposes we will regard transparency as being twofold, internal and external. Internal transparency is the extent by which the directorate and organisational branches of the Riksbank coordinate both their informational processing and the output produced by each consecutive branch. In our case, the relevant divisions are those of the Executive Board and the monetary policy office. We expect the relevant materials produced by these two units to be in conformity to one another. If this were not the case, and there is a biased tendency towards divergences in communication, we might draw the conclusion that information is not being efficiently transmitted between them. This would be reason enough to criticize the level of internal transparency. In contrast, external transparency only deals with one active part, the Riksbank, and one passive, the public, when it comes to information-transmission. In other words, the burden of responsibility lies on the Riksbank properly communicating its knowledge, intentions and forecasts; which in turn is entirely subject to the level of internal communication.

Thus, our approach has no need for considering anything other than what is produced by the Riksbank itself. As already established, the level of transparency can be entirely extracted from the documentations of the Riksbank, emphasising the fact that external transparency is a visualisation of the internal. By taking this approach we eliminate the possibility of confusing or misrepresenting the effects and nature of transparency because unlike previous studies, conducted by authors such as Kohn and Sack (2004) etc., we do not measure communication by its effect on external variables, but study it as a factor itself.

The of data for our analysis is composed of two groups of documents. The first group is the *Minutes of the Executive Board's monetary policy meetings*, which give an account of the Exectuive Board's discussions regarding the new repo rates (Sveriges Riksbank, 2016b). The minutes are as of 2008 published six times a year, and prior to this on a semi-regular basis. These in turn lay the foundations of our second group of documents, the *Monetary Policy Reports*, which reflect the considerations taken by the Executive Board when deciding monetary policy conduct (Sveriges Riksbank, 2016c). The reports also include various types of forecasts, such as the expected course of inflation given the Executive Board's decisions on interest rate. The reports are as of 2008 published six times a year, and prior to this between three to four times a year. These two groups constitute a *corpus* each – groups of documents that are to be processed for analytical purposes.

The data processing is carried out by the NLTK package of the Python programming language, a natural language processing toolkit. It reads all the available text included in our corpus, or any specified path therein, and produces a varying range of quantitative data, depending on what we are looking for. Once we have created the corpus, and assigned it a value within the Python User Interface, we can start using the methods that are included in the NLTK package to derive valuable information from our documents. The methods used are those outlined in the following segments.

3.1 Dictionary-identification technique

The first method, called the 'dictionary technique', is characterized by its straightforward arithmetic composition. It is essentially a percentage value representing how much a document is made up of a certain term or terms. It is initially carried out by defining any *dictionary* D, such that,

$$D = \{word_1, word_2, word_3, \dots word_n\} \qquad (1)$$

where $word_i$ are a collection of words related to each other under the same topic, such that the dictionary D pertains to a specific category of interest. In our analysis, we have made use of a number of dictionaries, D_i , encompassing the categories[1] we deem necessary for capturing topics regarding monetary policy. Such topics may be defined after purposes, which in our case encapsulate common categories concerned with monetary policy.

[1] For the complete list of terms specified to each category, see *section 4 (Results)*.

Once the count of any dictionary has been carried out for a document, the sum of all dictionary-terms are represented as a share of the total number of words contained in the document,

$$S_d = \frac{(word_1 + word_2 + word_3 + \ldots + word_n)}{number\ of\ words\ in\ document\ \boldsymbol{d}} \quad (2)$$

When we have produced a percentage-value of the dictionary contained within the document, we need an estimate to compare it to similar ones. In our case, the similarities between the two documents S_{d_1} and S_{d_2} can be compared by simple division. By doing this we can evaluate what fraction the dictionary-share of one document holds of another,

$$\frac{S_{d_1}}{S_{d_2}},\ only\ when\ S_{d_1} < S_{d_2} \quad (3)$$

As we are looking for how similar one document is to another, the comparison requires us to divide them with each other. However, the results will be nonsensical if we were to carry out a division in which the numerator is larger than the denominator, yielding a number larger than one. Thus, we only divide the smaller number with the larger, producing a value representing how similar the document-shares are to one another. The procedure is carried out on documents from the same period in time, year and month of publication, of both the minutes and the reports to estimate similarity in their usage of certain terms.

A rather well-known example of when this method has been used is in Paul Tetlock's modelling of the stock market by categorizing newspaper headlines after different sets of attitudes, positive and negative, and thus representing market swings in an alternative way (Tetlock, 2007). Though this is a very naïve, but effective, way of estimating the similarities between two documents, there exist some ways of refining it of which one is the following.

3.2 Weighting words

Merely counting the words can produce misleading results. Each word is unique in that it brings its own meaning into differing contexts, having one implication in one sentence, and another in the next. This stresses the weight and importance we attribute to each and every word, given its role in a text. For example, extremely common words such as '*if*', '*the*' and '*so*' beg the question of how much qualitative substance they add to understanding the key points of a text. Due to their highly frequent appearance in any given document, we might conclude that such words bear an almost insignificant role in the analysis; such words are called 'stop-words'

(Leskovec et al., 2014). These are often removed when analysing a text, or given far less weight than other more important words. Such a method of reasoning will, in our case, lead to a better interpretation of the results, giving less weight to words that appear particularly frequently, and more weight to words that are rarely used, yet bear a significant role in the terminology of monetary politics.

The way we achieve the weight-classifications is by making use of a technique called 'term frequency-inverse document frequency' (TF.IDF), which uses a logarithmic procedure to measure the frequency of a word in a certain document, and adding its weight (or impact) accordingly,

$$tf.idf_{t,d} = (1 + \log f_{t,d}) \cdot \log\left(\frac{D}{df_t}\right) \qquad (4)$$

where $\boldsymbol{D}$ is the total number of documents in the corpus**,** $\boldsymbol{df_t}$ is the number of documents in which the term $\boldsymbol{t}$ is present and finally $\boldsymbol{f_{t,d}}$ being the frequency of term $\boldsymbol{t}$ in a document $\boldsymbol{d}$ (Bholat et al., 2015). In our case, the relevant terms $\boldsymbol{t}$ are the terms that we assign to the dictionaries D_1 , D_2 , D_3 and D_4. The resulting TF.IDF's will measure the weight of every word in our dictionaries, in all documents of our corpus. Thus, in order to more easily represent the results, we add all the respective document-specific weights of each word and illustrate the TF.IDF as the sum of all those segments. Using the term *grow* from dictionary D_1 as an example, the resulting index will be produced by the following procedure,

$$tf.idf_{grow} = tf.idf_{grow,d_1} + tf.idf_{grow,d_2} + \ldots + tf.idf_{grow,d_n} \qquad (5)$$

By doing this, we are able to crunch up our table and still maintain the key results, if not even clearer than before.

3.3 Vector Space Models

It is of essential value to us that we are able to properly evaluate, not only how certain groups of words pertain to a specific document, but also the similarity between different documents as a whole, such that we are able to fulfil the aim of this essay. Some highly intuitive methods of accomplishing this, with regards to visualisation, use what are called *Vector Space Models* (figure 1). The way to go about understanding this is to think about documents as the sum of a number of words used in different amounts[2].

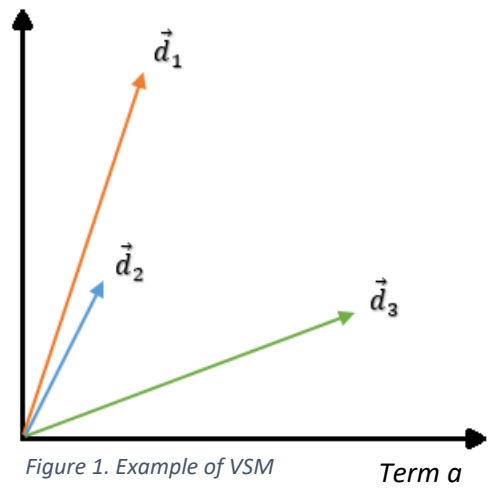

Figure 1. Example of VSM

Two documents might use the exact same terms, but as they focus on different things we would observe a variation of how these terms relate to each other. In *figure 1* two terms have been chosen, *term a* and *term b*, and shown are three documents $\overrightarrow{d_1}$, $\overrightarrow{d_2}$ and $\overrightarrow{d_3}$ using the two terms in varying quantities. With a simple representation like this we may gain certain insights, such that $\overrightarrow{d_2}$ and $\overrightarrow{d_3}$ seem to make use of *term b* in the same extent. The limitations of this is that we gain no information on how long the documents themselves are, giving us an indication as to whether two make us of the same term in the same extent. The approach is problematic precisely due to its limitations. For instance, if we were tasked with finding out which of the three documents in *figure 1* are the most alike we might get to very differing conclusions. For one, we might mention our previous observation and say that documents $\overrightarrow{d_2}$ and $\overrightarrow{d_3}$ must be very similar due their almost identical usage of the *term b*. On the other hand, we can also see that documents $\overrightarrow{d_1}$ and $\overrightarrow{d_2}$ can be considered similar, due to their almost identical usage of *term a*. Yet we would never equate the documents $\overrightarrow{d_1}$ and $\overrightarrow{d_3}$, proving a fallacy in our logics.

The issue that arises here stems from the limited methodology of merely counting the terms in each document, thus largely affecting any sensible comparisons between documents of

[2] This allows us to think of documents as vectors in an *n*-dimensional field - here *term a* and *term b* constitute a 2-dimensional field. An example of where this is used comes up in a paper by, among others, Antonina Kloptchenko et al., in which they measure the *Euclidian Distance* (length and value of each vector) of several financial reports and compare them accordingly (Kloptchenko et al., 2004).

differing length. For instance, if $\overrightarrow{d_1}$ was of greater length than $\overrightarrow{d_2}$ there would logically be more opportunity for $\overrightarrow{d_1}$ to include *term b* in its vocabulary, diminishing its potential likeness to $\overrightarrow{d_2}$, based on only *term a*. Thus, a more efficient way of measuring likeness between documents is through the usage of ratios between the terms in question. This can be deduced from the angle that is produced between two vectors, and can be calculated by the *Cosine Similarity* (CS) formula,

$$\cos\theta = \frac{\overrightarrow{d_i} \cdot \overrightarrow{d_j}}{\|\overrightarrow{d_i}\|\|\overrightarrow{d_j}\|} \qquad (6)$$

where $\overrightarrow{d_1} \cdot \overrightarrow{d_2}$ is the dot product between vectors $\overrightarrow{d_i}$ and $\overrightarrow{d_j}$, and $\|\overrightarrow{d_i}\|$ is the length of vector $\overrightarrow{d_i}$. The angle θ represents the degree by which two documents differ in their usage of some terms. As the angle grows, the documents are resemble each other less, and $\cos\theta$ approaches zero, and as the angle shrinks, the documents resemble each other more, and $\cos\theta$ approaches one. An example of CS being used is in an industry-arrangement analysis conducted by Hoberg and Philips (2010), in which they argue that their results provide a more satisfactory classification of businesses with regards to product-substitutability.

However, CS should not be taken for granted as a flawless solution to the previous mentioned dilemmas. The fact of the matter is that there are no clear-cut, systematic ways of writing reports and condensing ideas into text. In the end, the way in which we gather written information is ultimately shaped by the author of the document and their use of language, priority of subjects, values, level of professionalism, etc. Hence, we must constantly be aware of the impacts of inescapable obstacles in our analytical approach, such as the use of *synonymy* (using different words in reference of one thing) and *polysemy* (using a word that can have different meanings) (Bholat et al., 2015). The way in which we try and avoid these complications is by conducting our search through the documents with words that have an explicit meaning in the context of their respective documents. We have defined our dictionaries D_1 , D_2 , D_3 and D_4 in such a way as to try and minimize the mingling of ideas and contexts. By doing so, we marginalize the errors that can occur, so that we may make valuable use of methods such as Vector Space Models.

4. Results

The following sets of data were produced by the execution of Python directives by code, which carried out the methodological approaches we defined in *section 3*. They were extracted from two types of documents, the *Minutes of the Executive Board's Monetary Policy Meetings* and the *Monetary Policy Reports*, taken from the Riksbank's website. The timeline ranges from the year 1999 to the latest available documents of 2016. From 2008 onwards, both the minutes and the monetary policy reports have been coordinated to publication of six times a year. In the years 2006-2007, publication occurred only three times a year. From 1999-2005 the minutes had no standardized publication scheme, and thus the publications per year differed in numbers, unlike the monetary policy reports which were published four times a year during this period – once every quartile. When conducting our analysis, we have matched the monetary policy reports to their respective minutes, such that we end up with four pairs in the years 1999-2005, three pairs in 2006-2007, and six pairs in 2008-2016.

The latest report for 2016, i.e. that of period six, has yet to be published. Also, although notes for the minutes of the second period of 2016 have been published, no full excerpt is as of yet available, and thus the data is missing for that specific instance.

The terms we have chosen for the topic-specific dictionaries are as follows:

$$\mathbf{1}.\ D_1 = \{expect,\ forecast,\ future\}$$

$$\mathbf{2}.\ D_2 = \{appreciat,\ depreciat,\ exchange\}$$

$$\mathbf{3}.\ D_3 = \{employ,\ GDP,\ inflat,\ labour,\ pric,\ stabil\}$$

$$\mathbf{4}.\ D_4 = \{asset,\ bond,\ debt,\ risk, hous\}$$

The idea behind this system of classification is to try and encompass the many topics that are of importance to monetary policy evaluations. D_1 represents the discussions concerning the future expectations of the economy, D_2 is representative of global influence on the domestic economy, D_3 deals with the judgement of the current state and where we stand, and D_4 contains terms of matters that are of importance for economic evaluation in today's state.

In writing, terms can vary depending on the grammatical composition of the sentence. Thus, the terms will be cut off to their grammatical stems, for the purpose of targeting all variations of a term. An example being '*inflat*' which can be altered to produce '*inflation*', '*inflate*' and '*inflated*'. The term that sticks out is '*hous*' which is the shortened version of 'house price' or 'housing prices'. Due to the bending of the word 'house' to 'housing', 'houses', etc. we chose

to eliminate the word ‘price’ from ‘house prices’ and only search for ‘house’ by its grammatical stem ‘hous’. NLTK has no trouble finding parts of a word, as it only searches for the appointed characters, and does not equate them to singular words by themselves.

4.1 Results of the Dictionary-Identification Technique

The following tables are the derived from the calculated results of analysing our corpus according to the dictionary technique[3]. In order to better visualise the movements of the results through time, we have chosen to normalize them according to the average of each period of publication. Certain periods of interest have been colour-filled to indicate major events, or points of drastic changes. Such events are the financial crisis of 2007 which has been marked with dark-grey, so that we may compare the results both prior and post the crisis. We have also marked the years 2003 and 2006 with a line-break, to indicate years in which the Governor of Sveriges Riksbank has changed, going from Urban Bäckström to Lars Heikensten in 2003, and from Heikensten to the current Governor Stefan Ingves in 2006 (Sveriges Riksbank, 2011). The light shade of grey indicate periods of interest, where the likeness between the documents seem to have moved in an uncommonly unstable fashion, peaked or bottomed, compared to all other instances in time.

[3] For the raw and non-normalized results, refer to the *Appendix*.

	Forecasts, Expectations	Normalized values of the Dictionary-Identification Technique					
		1	**2**	**3**	**4**	**5**	**6**
Stefan Ingves	**2016**	0.29	-	1.27	1.14	0.89	-
	2015	1.17	1.05	1.07	1.02	1.28	1.16
	2014	0.82	1.03	0.97	1.14	0.95	1.04
	2013	1.15	0.99	0.93	0.90	0.85	0.90
	2012	0.88	1.05	0.78	0.94	0.95	0.80
	2011	1.11	0.84	0.90	0.99	0.89	1.10
	2010	1.20	0.80	1.22	1.02	0.88	1.19
	2009	0.81	0.90	0.94	0.76	1.07	0.93
	2008	1.01	1.17	1.00	1.05	1.23	0.89
	2007	1.19	1.01	1.22	-	-	-
	2006	1.26	0.94	1.03	-	-	-
Lars Heikensten	**2005**	0.77	0.98	1.13	0.96	-	-
	2004	1.15	1.15	0.78	0.83	-	-
	2003	0.84	1.08	0.99	1.11	-	-
Urban Bäckström	**2002**	1.32	1.05	1.01	1.08	-	-
	2001	0.90	1.09	0.89	1.10	-	-
	2000	1.24	0.95	1.12	1.02	-	-
	1999	0.89	0.93	0.73	0.95	-	-
	Mean	*0.76*	*0.84*	*0.78*	*0.86*	*0.77*	*0.82*

Table 1

	Exchange Rate	Normalized values of the Dictionary-Identification Technique					
		1	**2**	**3**	**4**	**5**	**6**
Stefan Ingves	**2016**	0.43	-	0.90	0.97	1.17	-
	2015	0.82	1.25	1.51	1.22	1.02	0.63
	2014	0.34	0.41	0.93	0.57	0.48	1.49
	2013	1.04	0.57	0.43	1.17	1.72	0.73
	2012	0.79	1.29	0.35	0.65	0.78	1.63
	2011	0.74	0.34	0.88	0.23	0.45	1.59
	2010	1.40	0.55	0.15	0.96	0.81	0.15
	2009	1.38	1.09	1.41	1.35	0.82	0.97
	2008	1.62	-	1.40	0.93	1.76	0.81
	2007	1.49	1.25	1.49	-	-	-
	2006	1.58	1.16	1.26	-	-	-
Lars Heikensten	**2005**	1.49	1.47	1.20	1.61	-	-
	2004	1.56	0.63	1.00	0.57	-	-
	2003	0.92	1.15	1.31	1.06	-	-
Urban Bäckström	**2002**	0.98	1.35	1.48	0.75	-	-
	2001	0.57	1.31	0.96	1.38	-	-
	2000	0.25	1.46	0.48	1.17	-	-
	1999	0.58	0.72	0.87	1.40	-	-
	Mean	*0.60*	*0.68*	*0.64*	*0.51*	*0.44*	*0.55*

Table 2

	Current Conditions	Normalized values of the Dictionary-Identification Technique					
		1	2	3	4	5	6
Stefan Ingves	**2016**	1.04	-	1.05	1.19	0.88	-
	2015	1.04	1.03	0.87	1.13	1.17	1.09
	2014	1.08	1.06	0.99	1.01	0.81	1.04
	2013	1.03	1.00	0.93	1.01	0.66	1.08
	2012	1.11	1.15	1.22	1.12	1.18	1.00
	2011	0.97	1.11	0.74	0.93	1.20	0.95
	2010	0.96	0.84	1.15	0.70	1.10	0.61
	2009	1.10	0.92	0.91	0.83	1.02	1.09
	2008	1.01	1.07	1.05	1.15	0.99	1.14
	2007	1.03	1.13	1.00	-	-	-
	2006	0.65	1.02	0.99	-	-	-
Lars Heikensten	**2005**	1.06	0.96	0.73	0.78	-	-
	2004	1.07	0.64	0.84	1.06	-	-
	2003	0.73	1.10	1.21	1.21	-	-
Urban Bäckström	**2002**	1.07	0.86	0.96	0.73	-	-
	2001	0.88	1.04	1.15	1.07	-	-
	2000	1.08	1.03	1.09	1.04	-	-
	1999	1.11	1.05	1.12	1.05	-	-
	Mean	*0.85*	*0.86*	*0.82*	*0.78*	*0.83*	*0.88*

Table 3

	Important Matters	Normalized values of the Dictionary-Identification Technique					
		1	2	3	4	5	6
Stefan Ingves	**2016**	1.11	-	1.37	0.75	1.13	-
	2015	0.89	1.47	1.54	1.25	1.22	1.28
	2014	1.45	0.86	0.81	1.39	1.22	1.26
	2013	1.85	0.55	0.82	1.01	0.65	0.87
	2012	1.03	1.38	1.42	1.27	0.41	0.47
	2011	1.40	1.36	1.05	1.36	1.20	0.89
	2010	1.41	1.01	1.18	1.28	1.06	1.09
	2009	1.30	0.80	1.64	1.03	1.08	1.11
	2008	1.51	0.70	1.25	1.40	1.04	1.02
	2007	0.42	0.88	0.63	-	-	-
	2006	0.69	0.71	0.54	-	-	-
Lars Heikensten	**2005**	0.48	0.62	0.83	0.45	-	-
	2004	0.59	0.53	0.81	0.74	-	-
	2003	0.46	1.43	0.70	0.53	-	-
Urban Bäckström	**2002**	0.61	1.43	1.10	0.52	-	-
	2001	1.18	1.20	0.88	1.48	-	-
	2000	0.97	0.51	0.67	0.34	-	-
	1999	0.64	1.56	0.76	1.20	-	-
	Mean	*0.50*	*0.63*	*0.60*	*0.66*	*0.77*	*0.76*

Table 4

There are given instances in the tables in which the data appears to be missing. This is due to one of three reasons:

1. No documents have been published, as is the case with *Policy2016_6* which is due to come out in 2017, or that some documentation is still under processing and is yet to be published in its complete form, as is the case with *Minute2016_2*, originally published in January of 2016.

2. The number of documents published by the Riksbank throughout the year has changed in the period 1999-2016. As mentioned earlier, from 2008 and onwards six reports, of each respective category, are published each year, once every other month. During the period 2006-2007 only three reports of each document are published each year, once every four months. Lastly, from 1999-2005 there was no established structure as to when the documents were published, with respect to each other. It was often the case that many Monetary Policy Reports could discuss a single Minute of the Executive Board. As such, we compile the number of both categories down to four due to the Minutes being published once every quartile, and used the latest Monetary Policy Report discussing a given Minute to correspond to all of them, leaving us with four Reports a year as well.

3. Problems with compiling data afflicted by the Python script used, being unable to extract the necessary information for analysis, or the case that the words we were searching for did not exist in the files. This is the case of the documents from the period 2008 (2), dictionary D_2.

4.2 Results of the TF.IDF

In the following table, and as proposed in *section 3.2*, all individual TF.IDF's have been summed into their respective dictionaries, in order for the difference in weight, and importance, to be better understood.

D1			D2			D3			D4	
expect	10.51		*appreciat*	76.33		*employ*	16.90		*asset*	85.06
forecast	5.11		*depreciat*	107.98		*GDP*	24.07		*bond*	60.45
future	17.32		*exchange*	19.50		*inflat*	11.06		*debt*	87.96
						labour	16.60		*risk*	8.22
						pric	4.97		*hous*	13.47
						stabil	36.96			

Table 5. TF.IDF values of all dictionary terms.

The highlighted terms in the table are of particular interest. Darker shades of grey indicate words that bare higher significance in the context of where they are used. Such words are 'appreciation' or 'depreciation'. Lighter shades of grey indicate words that are used so regularly that they bare next to no significant impact on the document, when they are used. These values will provide greater insight into the performance of communication, as estimated in *sections 4.2* & *4.3*.

4.3 Results of the Cosine Similarities

The following results describe the cosine similarity between the minutes and the monetary policy reports, by the method outlined in *section 3.3*.

	1999	2000	2001	2002	2003	2004	2005	2006	2007
1	0.83	0.92	0.91	0.86	0.84	0.88	0.85	0.67	0.89
2	0.88	0.85	0.92	0.85	0.94	0.74	0.69	0.99	0.90
3	0.91	0.89	0.91	0.83	0.91	0.73	0.71	0.83	0.93
4	0.70	0.71	0.98	0.82	0.93	0.86	0.72		

	2008	2009	2010	2011	2012	2013	2014	2015	2016
1	0.86	0.84	0.89	0.91	0.76	0.84	0.85	0.90	0.76
2	0.83	0.84	0.82	0.89	0.79	0.80	0.86	0.96	-
3	0.93	0.88	0.96	0.82	0.83	0.85	0.85	0.95	0.85
4	0.88	0.90	0.92	0.77	0.75	0.85	0.77	0.95	0.90
5	0.90	0.88	0.94	0.82	0.83	0.75	0.88	0.88	0.93
6	0.88	0.86	0.86	0.79	0.72	0.86	0.91	0.88	-

Table 6. Cosine similarity between every minute and report, for all time periods.

The years of 2003, 2006 and 2007 are marked with a darker shade of grey, indicating periods of significant change, due to the reasons mentioned earlier. The values indicated by the lightest

shade of grey are periods of particularly low similarity, specifically 75 percent and below. The possible reasons for their values will be discussed in the following section.

5. Discussion

The data we have gathered thus far might give rise to differing conclusions, depending on how we approach the set of results and in what manner we define our questions. For our purposes, we are chiefly interested in what extent our two corpora, the *Minutes of the Executive Board's Monetary Policy Meetings* and the *Monetary Policy Reports*, deviate from one another. Our analytical approaches have given us a variety of methods for conducting, and interpreting, this difference in.

First, the dictionary technique has helped us identify the similarities in usage of specific topic-related terms, during different periods in time. By this, we are able to pinpoint certain movements, in the similarities, of interest during different time intervals, matching them with one another and draw possible conclusions of 'favouritism' of one group of terms over another. Further, it enables us to correlate certain swings in similarity during times of great change, such as that of a Governor of the Riksbank, or major market movements such as the financial crisis of 2007-2008.

Second, recognizing the insufficient approach of the dictionary technique, we take certain steps of approaching the actual context of the documents, instead of measuring it at face value. The TF.IDF allows us to measure the importance – weight – of all the terms in our dictionaries by comparing them in the regularity of their appearance, thus providing us with a certain measure of their impact, whenever they might appear. This additional analytical tool is of crucial importance when looking over the results of the dictionary technique. This is due to the fact that as we might observe certain dictionaries bearing greater importance and impact than others, we necessarily assume an exceptional vitality in high similarities between the two types of documents concerning the terms related to it. A proper system of transparency would undoubtedly require matters of greater importance to be communicated in as correct a way as possible.

Finally, the technical approach of vector space modelling provides us with a deeper insight into how the documents as a whole differ from each other, with no focus on any specific terms. This is important because, besides being a means of validating our previous results, it can be used as a way of modelling the movements of transparency through time. Also of great importance

is to examine the values during times of change, such as the financial crisis, as to add greater insights into how big events impact communication – both during the time of the event, and as a lasting effect for the years to come. By examining the average movements of these values, we may even give preliminary forecasts of how the development of communication appears to be moving in the future, with regards to historical movements.

5.1 Communication before and after the financial crisis

It escapes nobody that the financial crisis of 2007-2008 affected almost every corner of the global economy. What is of interest to us is to see whether it had any impact on the communicatory conformity of the Riksbank. In order to do this, we need to examine the performance of transparency before, during, and after the financial crisis. If there was an effect, we expect to see some form of dramatic plunges in dictionary-similarities in the periods 2007-2008, and a lagging effect following it for the years after. However, it is also possible that some dictionaries would have improved in resemblance, with regards to the crisis. If that is the case, we expect dictionary D_4 and possibly D_3 to be the ones with the increase, as the former deals with risk, debt and housing prices, and the latter deals with GDP and unemployment; all very crucial topics of discussion during and after the crash.

In order to examine the movements of term-similarities, we need to make use of the results derived from the dictionary technique. Apart from the normalization conducted in *section 4*, *figure 2* has been put together, as the average movement of all values, to map out all the dictionaries and their respective development through time.

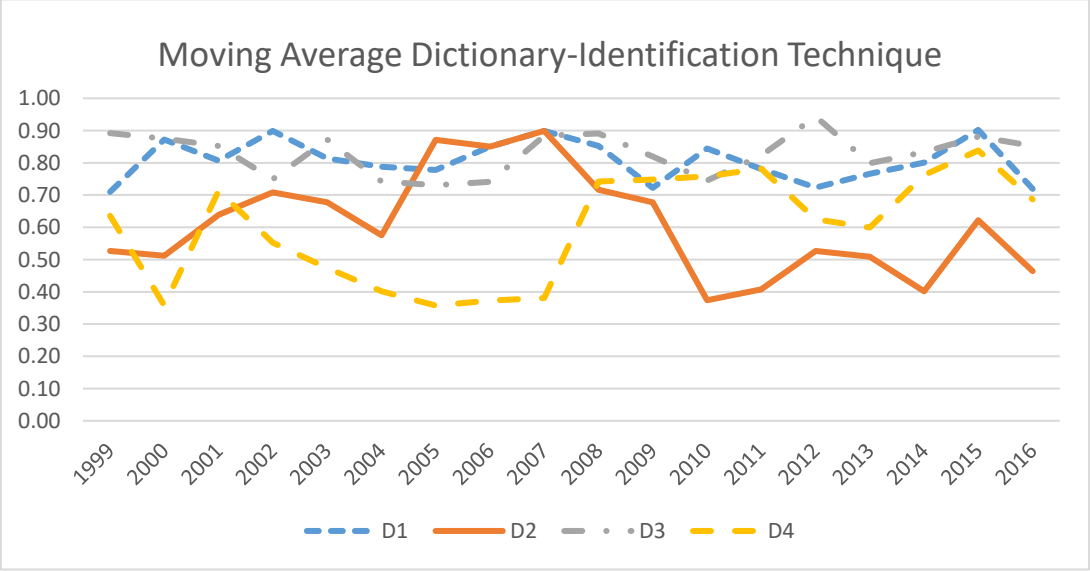

Figure 2. Moving average of the Dictionary Technique data for all dictionaries.

As can be observed, there are a number of interesting movements happening. A breaking point in the graph seems to be in the year 2007, for all dictionaries. As shown in the table 1 and table 2, up until 2007 the dictionaries D_1 and D_2 seem to have moved in a rather steady level in-between 80 and 90 percentage points, from 2005 for D_2 and all the way back to 2000 for D_1. After the peak in 2007 they decline and do not regain their upwards trend until some years later, 2010 for D_1 and 2011 for D_2. This might very well be an indication of the financial crisis having affected the way in which the Riksbank handles communication, seeing a definitive decline in document-similarity in areas concerning future expectations and appreciations/depreciations.

At the same time we are witnessing the exact opposite reaction for the remaining two dictionaries. Dictionary D_4 shows a remarkable increase in similarity, from around 40 percentage points to 75 percentage points, as seen in table 4; this after having experienced a steady decrease in similarity since 2001, as shown by the graph. In an interesting move, dictionary D_3 seems to predicate the upwards movement of D_4 by a whole year, in 2006, going up from around 70 to 90 percentage points and remains there for the duration of the financial crisis, 2007 to 2008.

It stands thus clear that the financial crisis had an impact on the transparency of the Riksbank, in quite different ways. Terms that are directly linked to the central aspects of the financial crisis, such as ‘housing bubble’, ‘debt’ and ‘unemployment’ seem to have gained tremendous traction during this period. It is evident that the organisation of the Riksbank recognised the importance of properly communicating the insights it had, regarding the issues immediately concerned with the crisis. This event might also have had positive long-term effects on the Riksbank, as we can see the average movements of dictionary D_4 remain higher after the crisis than before, indicating a prolonged interest in the subjects concerning its terms. Unfortunately, this seems to have come at the cost of a decreased focus on the other dictionaries, D_1 and D_2, which are made up of terms not immediately linked to the crisis. Dictionary D_1 seems to have recovered rather quickly, but D_2 is as of yet affected, currently hovering at an average level below any year previous to the crisis; although the graph does seem to indicate some manner of an upward-trend.

To further develop the significance of these movements, we bring into account the calculated weights of the TF.IDF’s. With these in mind, we may better understand the reason behind the movements, or point out potentially critical flaws in terms of communication. To better

visualize the values, *figure 3* represents the summed TF.IDF weights of all terms for all periods, in both documents.

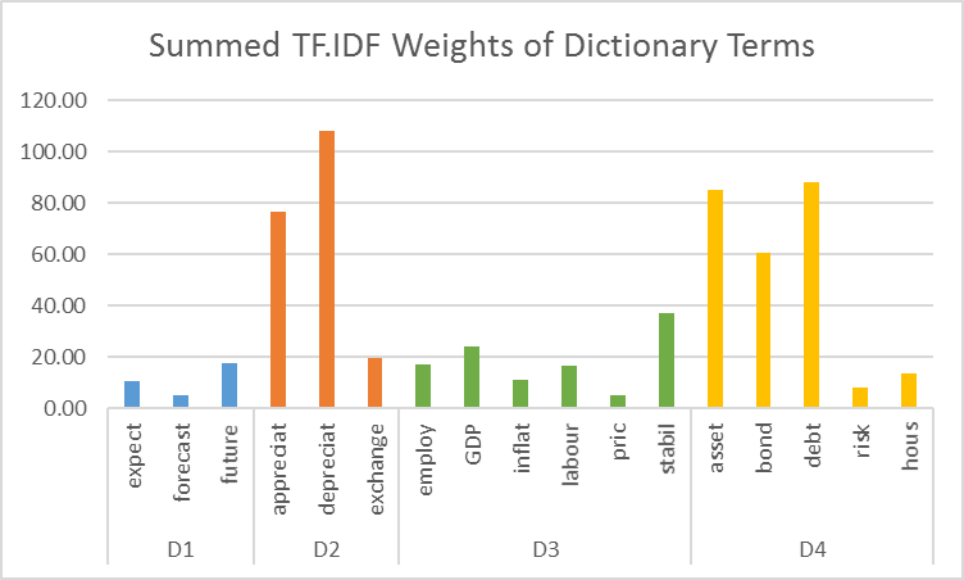

Figure 3. Diagrams of the summer TF.IDF weights for all terms.

As expected, we find terms such as 'risk' and 'house price' to bear next to insignificant weight, indicating their regular usage most likely during financial crisis. The remaining terms in D_4 tell another story. Taking into account the extremely low value of document-similarity prior to the financial crisis, the high weight of the remaining terms in D_4 would indicate a serious flaw in properly communicating topics of debt and the question of leverage and risk. However, the upward-movement of D_4 in figure 2 marks the improvement of communication in this area, meaning that the financial crisis served the good purpose of attracting more focus to these topics – a significant improvement of communication.

On the other hand, the work done with properly communicating dictionary D_2 seems to have taken a critically negative turn. As can be seen, the words 'appreciation' and 'depreciation' hold amongst the highest level of weight in the whole corpus, indicating the importance of coordinating their communication. What we however observe is, as previously mentioned, a decreased level of integration between the minutes and the monetary policy reports regarding these terms, after the financial crisis. Given their extremely high weight on any document of which they are part, we may conclude that this is an impairing turn of events on behalf of the Riksbank. The case of transparency regarding the issues of the exchange rate (imported inflation) is at an intolerable level, and it may even be called into question of whether the Riksbank is being transparent in this area, or not. What stands clear is that vast improvement

in the communication in this field needs to take place, and what can opportunistically be seen as an upwards-trend from 2011-2016 may be deemed as a way of resolving this issue.

5.2 Communication under different Governors of the Riksbank

The possibility of a general performance test in the similarities between the two types of documents is possible due to our method of vector space modelling. This approach enables us get an overview of the performance for the complete time-horizon 1999-2016, and make possible connections to specific events in time. Such a recurring event, that spans segments of the timeframe, is the replacement of the Governor of the Riksbank. In *figure 4* we find the moving average of all cosine similarity values, as well as a fitted polynomial curve to indicate long-term movements.

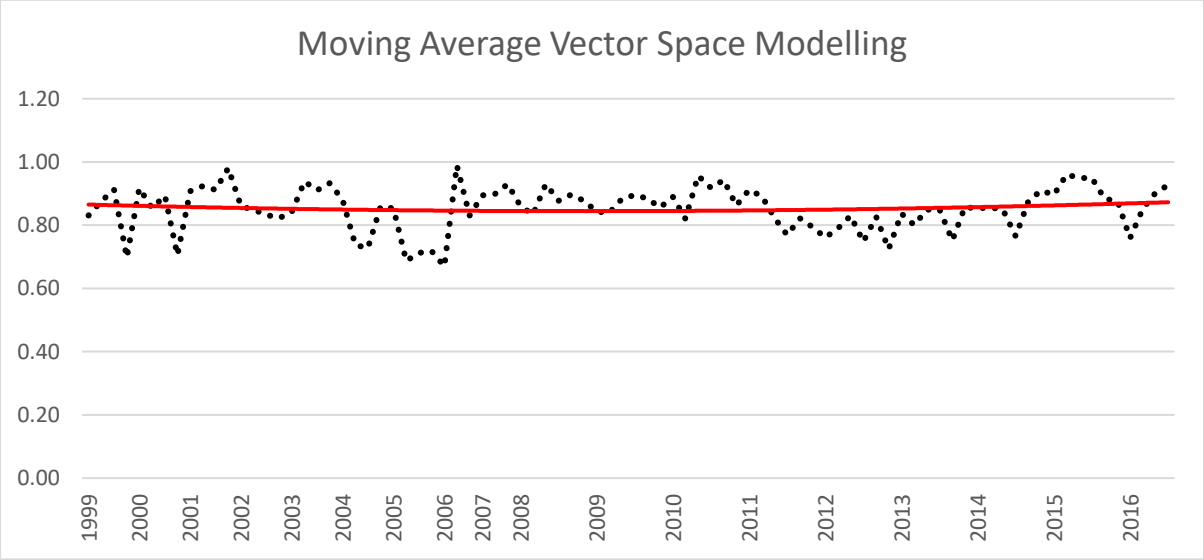

Figure 4. Moving average of the Cosine Similarity values for all periods.

In our time-frame, there have been three Governors i.e. two periods of replacement. Urban Bäckström served as Governor from 1994-2002, handing over the position to Heikensten in the beginning of 2003. Heikensten remained in office as Governor for only two years, until the end of 2005, handing it over to our current Governor, Stefan Ingves, in 2006. Thus we have three periods of interest: 1999-2002, 2003-2005, 2006-2016.

The most remarkable period of low cosine similarity seems to take place between 2004-2006, in which there seems to be a steady downward movement; in table 6 we find the lowest values falling from 74 points in 2004, all the way down to 67 points in 2006, a remarkably low value in comparison to all others. A possibility might be linked to Heikensten running as Governor during this period, although there is nothing that indicates his administration is responsible for

worsening the Riksbank's ability to communicate. What happens immediately after 2006 is the extreme jump in similarity, going from an average of 70 points to around 95 points in 2007. This can possibly be linked to the newly installed Governor Stefan Ingves, although it is most likely linked to the increased responsibility of managing the Riksbank's activities during the wake of the financial crisis.

An interesting observation occurs in the polynomial fit, which seems to indicate a U-shaped curvature. This is both indicative of a poorer handling of communication in the middle time-frame, but also a sign of improvement and a possible forecast of greater transparency in the future. Evidently, the role of Governor seems to have minor influence upon the rate of correspondences between the minutes and the reports. Although some remarks may be made according to our tables, influence seems reserved to major economic events instead.

6. Conclusion

As the role of communication has grown over time, the present disposition of the central banks is to advocate for a more transparent means of operation. This means that they take great concern with the internal rate of transparency, within the organisation. This has shown to largely affect the public's expectations of central bank actions, and is of vital concern when designing monetary policy. We have shown how the level of transparency of the Riksbank is significantly affected by major events, such as the financial crisis or, to some extent, the switching of Governor. As has been shown, there is major work left to be done and priority ought to be set on diversifying communication when writing reports. This is important as to not be swayed by major events surrounding the operations of the Riksbank, thus altering the qualitative level by which the reports are being produced.

References

Apel, M. and Vredin, A. (2007). "Monetary-Policy Communication: The Experience of the Swedish Riksbank". *Czech Journal of Economics and Finance,* **57**: pp. 499-520.

Bernanke, B. S. (2004). "Remarks at the Meetings of the American Economic Association, San Diego, California, January 3, 2004". *Fedspeak.* Available at https://www.federalreserve.gov/boarddocs/speeches/2004/200401032/default.htm [accessed 23/12/2016, at 17:15].

Bholat, D., Hansen, S., Santos, P., Schonhardt-Bailey, C. (2015). "Text mining for central banks". *Centre for Central Banking Studies: Handbook,* **33**: p. 9.

Bird, et al. (2015). "Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit". *NLTK.* Available at http://www.nltk.org/book/ch01.html [accessed 29/12/2016, at 14:10].

Blinder, S. A., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D.J. (2008). "Central Bank Communication and Monetary Policy: A survey of theory and evidence". *European Central Bank,* **898**: p. 7.

Connolly, E. and Kohler, M. (2004). "News and Interest Rate Expectations: A Study of Six central Banks". *The Future of Inflation Targeting,* Sydney: Reserve Bank of Australia, pp. 108-34.

Dincer, N. N. and Eichengreen, B. (2013). "Central Bank Transparency and Independence". *The Bank of Korea – Working/Dicussion Papers.*

Ehrmann, M. and Fratzscher, M. (2005). "How should central banks communicate?". *European Central Bank working paper series,* **557**.

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

Faust, J. and Svensson, L. E. O. (2001). "Transparency and Credivility: Monetary Policy with Unobservable Goals". *International Economic Review,* **42**: pp. 369-397.

Goodfriend, Marvin (1991). "Interest Rates and the Conduct of Monetary Policy". *Carnegie-Rochester Conference Series on Public policy,* **34**: pp. 7-30.

Hoberg, G., Philips, G. (2010). "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis". *The Review of Financial Studies: Oxford Journals,* **23**: pp. 3773-3811.

Issing, O. (2005). "Communication, Transarency, Accountability – Monetary Policy in the Twenty-First Century". *Federal Reserve Bank of St. Louis Review,* **87**: pp. 65-83.

Jansen, D. and De Haanm, J. (2005). "Talking Heads: The Effects ifi ECB Statements on the Euro-Dollar Exchange Rate." *Journal of International Money and Finance,* **24** (2): pp. 343-361.

Kang, Y. et al. (2013). “Central Bank Communication Policy: A Comparative Study”. *School of International and Public Affairs, Colombia University*. New York.

Kindler, E. and Krivy, I. (2011). “Object-oriented simulation of systems with sophisticated control”. *International Journal of General Systems*, **40** (3): pp. 313-343.

Kloptchenko, A., Magnusson, C., Back, B., Visa, A., Vanharanta, H. (2004). ”Mining Textual Contents of Financial Reports”. *The International Journal of Digital Accounting research*, **4**: pp. 1-29.

Kohn, D. L. and Sack, B. (2004). “Central Bank Talk: Does it Matter and Why?”. *Macroeconomics, Monetary Policy, and Financial Stability Conference*, Ottowa: Canada.

Leskovec, J., Rajaraman, A., Ullman, J. D. (2014). ”Chapter 1: Data Mining”. *Mining of Massive Datasets*, p. 8.

Mishkin, F. S. (2004). “Can Central Bank Transparency Go Too Far?”. *The National Bureau of Economic Research, Reserve Bank of Australia*, Conference-2004: pp. 48-65.

Mishkin, F. S. (2014). “Central Banking After The Crisis”. Chapter 2 in *Macroeconomic and Financial Stability: challenges for Monetary Policy*, **19**: pp. 23-59.

Poole, W. (2001). “Expectations”. *Federal Reserve Bank of St. Louis Revview*, **93** (1): pp. 1-10.

Reeves, R. and Sawicki, M. (2007). “Do Financial Markets React to Bank of England Communication?”. *European Journal of Political Economy*, **23** (1): pp. 207-227.

Sveriges Riksbank (2011). “Former governors of the Riksbank”. *The Riksbank*. Available at http://www.riksbank.se/en/The-Riksbank/History/Tidigare-riksbankschefer-finns-enbart-pa-sv-sida/
[accessed 2/1/2017, at 15:30].

Sveriges Riksbank (2016a). “The Riksbank’s communication policy”. *The General Secretariat*

Sveriges Riksbank (2016b). “Minutes of the Executive Board’s monetary policy meetings”. *Press & Published*. Available at http://www.riksbank.se/en/Press-and-published/Minutes-of-the-Executive-Boards-monetary-policy-meetings/?all=1
[accessed 1/1/2017, at 19:40].

Sveriges Riksbank (2016c). “Monetary Policy Reports”. *Press & Published*. Available at http://www.riksbank.se/en/Press-and-published/Published-from-the-Riksbank/Monetary-policy/Monetary-Policy-Report/?all=1
[accessed 1/1/2017, at 19:40].

Tetlock, Paul C. (2007). “Giving Content to Investor Sentiment: The Role of Media in the Stock Market”. *The Journal of Finance*, **62**: pp. 1139-1168.

Woodford, M. (2001). “Monetary Policy in the Information Economy”. *National Bureau of Economic Research*, p. 12.

Woodford, M. (2004). “Inflation targeting and Optimal Monetary Policy”. *Federal Reserve Bank of St. Louis Economic Review*, **86**(4): pp. 15-41.

Woodford, M. (2005). “Central-Bank Communication and Policy Effectiveness”. *The Greenspan Era: Lessons for the Future: A Symposium Sponsored by the Federal Reserve Bank of Kansas City*, Federal Reserve Bank of Kansas City, p. 399-474.

Woodford, M. (2012). “Method of Policy Accommodation at the Interest-Rate Lower Bound”. *Colombia University*.

Appendix

The following array of tables present the raw results of applying the Dictionary-Identification Technique on our corpus.

Forecasts, Expectations	Dictionary-Identification Technique					
	1	**2**	**3**	**4**	**5**	**6**
2016	0.22	-	0.99	0.98	0.69	-
2015	0.88	0.88	0.84	0.88	0.98	0.95
2014	0.62	0.86	0.76	0.98	0.73	0.85
2013	0.87	0.83	0.73	0.77	0.65	0.73
2012	0.66	0.88	0.61	0.81	0.73	0.65
2011	0.84	0.70	0.70	0.85	0.69	0.90
2010	0.91	0.67	0.96	0.88	0.68	0.98
2009	0.61	0.75	0.73	0.65	0.83	0.76
2008	0.77	0.99	0.78	0.91	0.95	0.72
2007	0.90	0.84	0.96	-	-	-
2006	0.95	0.79	0.81	-	-	-
2005	0.58	0.82	0.88	0.82	-	-
2004	0.87	0.96	0.61	0.71	-	-
2003	0.63	0.90	0.77	0.95	-	-
2002	1.00	0.88	0.79	0.93	-	-
2001	0.68	0.91	0.69	0.94	-	-
2000	0.94	0.80	0.88	0.87	-	-
1999	0.67	0.78	0.57	0.82	-	-

Exchange Rate	Dictionary-Identification Technique					
	1	**2**	**3**	**4**	**5**	**6**
2016	0.26	-	0.58	0.49	0.52	-
2015	0.49	0.85	0.97	0.62	0.45	0.35
2014	0.21	0.28	0.60	0.29	0.21	0.83
2013	0.63	0.39	0.27	0.60	0.76	0.40
2012	0.48	0.87	0.23	0.33	0.35	0.90
2011	0.45	0.23	0.56	0.12	0.20	0.88
2010	0.84	0.37	0.10	0.49	0.36	0.08
2009	0.83	0.74	0.91	0.69	0.37	0.54
2008	0.98	-	0.90	0.47	0.78	0.45
2007	0.90	0.84	0.96	-	-	-
2006	0.95	0.79	0.81	-	-	-
2005	0.90	1.00	0.77	0.82	-	-
2004	0.94	0.43	0.64	0.29	-	-
2003	0.55	0.78	0.84	0.54	-	-
2002	0.59	0.91	0.95	0.38	-	-
2001	0.34	0.89	0.62	0.71	-	-
2000	0.15	0.99	0.31	0.60	-	-
1999	0.35	0.49	0.56	0.71	-	-

Current Conditions	Dictionary-Identification Technique					
	1	**2**	**3**	**4**	**5**	**6**
2016	0.88	-	0.86	0.93	0.73	-
2015	0.88	0.88	0.71	0.88	0.98	0.96
2014	0.92	0.90	0.80	0.79	0.67	0.91
2013	0.87	0.86	0.76	0.79	0.55	0.95
2012	0.94	0.98	1.00	0.88	0.98	0.88
2011	0.82	0.95	0.60	0.73	1.00	0.83
2010	0.82	0.72	0.94	0.55	0.92	0.53
2009	0.93	0.79	0.74	0.65	0.85	0.96
2008	0.85	0.91	0.86	0.90	0.82	1.00
2007	0.87	0.97	0.82	-	-	-
2006	0.55	0.87	0.81	-	-	-
2005	0.90	0.82	0.60	0.61	-	-
2004	0.91	0.55	0.68	0.83	-	-
2003	0.62	0.94	0.99	0.94	-	-
2002	0.91	0.74	0.78	0.57	-	-
2001	0.75	0.89	0.94	0.83	-	-
2000	0.91	0.89	0.89	0.81	-	-
1999	0.94	0.90	0.91	0.82	-	-

Important Matters	Dictionary-Identification Technique					
	1	**2**	**3**	**4**	**5**	**6**
2016	0.56	-	0.82	0.49	0.87	-
2015	0.45	0.92	0.92	0.82	0.94	0.97
2014	0.73	0.54	0.48	0.92	0.94	0.95
2013	0.93	0.34	0.49	0.66	0.50	0.66
2012	0.52	0.87	0.85	0.84	0.32	0.35
2011	0.71	0.85	0.63	0.89	0.93	0.68
2010	0.71	0.64	0.71	0.84	0.82	0.83
2009	0.65	0.50	0.98	0.68	0.83	0.84
2008	0.76	0.44	0.75	0.92	0.81	0.77
2007	0.21	0.56	0.38	-	-	-
2006	0.35	0.44	0.33	-	-	-
2005	0.24	0.39	0.50	0.30	-	-
2004	0.30	0.33	0.49	0.49	-	-
2003	0.23	0.90	0.42	0.35	-	-
2002	0.31	0.90	0.66	0.34	-	-
2001	0.59	0.76	0.53	0.98	-	-
2000	0.49	0.32	0.40	0.23	-	-
1999	0.32	0.98	0.45	0.79	-	-