연세대학교 상경대학
경제연구소
Economic Research Institute
Yonsei University

서울시 서대문구 연세로 50
50 Yonsei-ro, Seodaemun-gu, Seoul, Korea
TEL: (+82-2) 2123-4065 FAX: (+82-2) 364-9149
E-mail: yeri4065@yonsei.ac.kr
http://yeri.yonsei.ac.kr/new

Deciphering Monetary Policy Committee Minutes with Text Mining Approach: A Case of South Korea

Youngjoon Lee
Yonsei University
Soohyon Kim
Bank of Korea
Ki Young Park
Yonsei University

October 2018
2018RWP-132

Economic Research Institute
Yonsei University

Deciphering Monetary Policy Committee Minutes with Text Mining Approach: A Case of South Korea*

Youngjoon Lee† Soohyon Kim‡ Ki Young Park§

October 10, 2018

Abstract

We quantify the Monetary Policy Committee (MPC) minutes of the Bank of Korea (BOK) using text mining approach. We propose a novel approach using a field-specific Korean dictionary and contiguous sequence of words (n-grams) to better capture the subtlety of central bank communication. We find that our lexicon-based indicators help explain the current and future BOK monetary policy decisions when considering an augmented Taylor rule, suggesting that they contain additional information beyond the currently available macroeconomic variables. Our indicators remarkably outperform English-based textual classifications, a media-based measure of economic policy uncertainty, and a data-based measure of macroeconomic uncertainty. Our empirical results also emphasize the importance of using a field-specific dictionary and the original Korean text.

JEL classification: E43, E52, E58

Keywords: monetary policy; text mining; central banking; Bank of Korea; Taylor rule

*The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of Korea. The usual disclaimers apply.

†yj.lee@yonsei.ac.kr, School of Business, Yonsei University.

‡soohyonkim@bok.or.kr, Bank of Korea.

§kypark@yonsei.ac.kr, School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, 120-749, South Korea.

1. Introduction

As the title of Gentzkow, Kelly, and Taddy (2017), "text as data," succinctly encapsulates, text is data in that it is useful sources of information. Meanwhile, it has been underutilized since it is difficult to quantify and interpret compared to numerical data. As computing power has increased, text mining analysis has evolved from a labor-intensive manual discipline to a sub-field of "big data" analysis. And, as Bholat, Hansen, Santos, and Schonhardt-Bailey (2015) point out, the computer-enabled approach of text mining can not only process more texts than any person can read but also extract information that might be missed by human readers.

While text mining has been more actively applied to other fields such as marketing and political sciences, it has not been a main tool for economic analysis. It is particularly the case for monetary policy and macroprudential policy. However, given the increased importance of transparency in conducting monetary policy, it is important to develop and apply appropriate tools to analyze central bank communication. Warsh (2014), a former Governor of the Federal Reserve Board, emphasizes the importance of the textual discourse as a potential source of additional information by saying "No surprise, Fed policymakers far more often reveal their differing judgments on economic variables in their discussion around the table than in their actual votes" And he highly evaluate the usefulness of text mining approach by saying "A more rigorous and constructive means of judging the effects of the Fed's new transcript policy can be found by evaluating the text of the transcripts themselves."

In this regard, we take text mining approach to extract the quantitative information on monetary policy decision making from 232,658 documents during the period of May 2005–December 2017.[1] By converting qualitative contents in the Bank of Korea MPC (Monetary Policy Committee) minutes into quantitative indicators, we measure the sentiment and tone of the minutes and examine if the MPC communication conveys additional information that are not included in the available macroeconomic data. We find that our lexicon-based indicators help explain the current and future monetary policy when considering an augment Taylor rule. In addition, our indicators significantly outperform English text-based indicators, a media-based measure of economic policy uncertainty measure based on Baker, Bloom, and Davis (2016) and a measure of macroeconomic uncertainty measure developed by Jurado, Ludvigson, and Ng (2015). By comparing with various measures, we provide a guidance on the direction of future research in this field. Our study clearly shows the importance of using a field-specific dictionary and the original Korean text (not Korean-to-English text).

[1] We use 151 minutes of the MPC (Monetary Policy Committee), 206,223 news articles related to interest rates, and 26,284 bond analyst reports. We provide a more detailed explanation on our data in Section 3.

We make unique contributions in several aspects. First, to our knowledge, this is the first study that applies the sentiment analysis to monetary policy decision making of the BOK (Bank of Korea) and demonstrates that lexicon-based indicators have ample information on monetary policy beyond information in macroeconomic variables. Our result suggests several venues for future research. For example, one may interpret our indicators as a latent variable of the BOK policy rate and feed them into the standard VAR or DSGE models that analyze the effect of monetary policy. Or our indicators can be used for evaluating the effectiveness of the BOK's communication on its future direction of monetary policy. Given that central bank communication has emerged as an important tool for central banks to manage public expectations on inflation and economic activities, it is important to analyze what kinds of information the BOK documents convey. In this regard, our study demonstrates the usefulness of text mining approach to tasks related to central banking.

Second, in terms of methodology, we adopt the several advanced techniques in this field. Since it is not easy to determine the tone or sentiment of a single word (uni-gram) or a bi-gram phrase that combines positive and negative words like "lower unemployment" or "sluggish recovery," we use n-grams. With n-grams (from 1-gram to 5-gram) as features, we consider contexts and better capture the subtlety of central bank communication. In order to determine the polarity (hawkish/neutral/dovish in our case) of these n-grams, we develop two kinds of sentiment indicators based on two contrasting but complementing methods. One is the market approach and the other is the lexical approach. One advantage of the market approach is that it does not rely on the researchers' subjective selection of seed words and uses only the market information. However, since it decides the polarity of a word (n-gram in our case) based on its statistical association with market information (e.g., stock returns), indicators based on this approach may naturally produce statistically significant outcomes when we examine the market impact of indicators. In contrast, the lexcial approach decides the polarity based on the proximity to the pre-determined seed words. It is clear that its performance depends on choices of seed words. To reduce the problem of researchers' discretion over seed words, we use a state-of-the-art domain-specific sentiment induction algorithm called the SentProp framework by Hamilton, Clark, Leskovec, and Jurafsky (2016). Further, to evaluate the accuracy of our lexicon classification, we use documents that are not used in building our lexicons. We manually label 2,341 sentences of the introductory statements from the BOK Governor's news conference and perform an out-of-sample test. We confirm that the accuracy is quite high.

Third, we use our own natural language processing (NLP) tool called eKoNLPy (Korean NLP Python Library for Economic Analysis) to address the difficulties associated with the Korean language such as field-specific foreign words, irregular conjugation of verbs and

adjectives, and others. eKoNLPy is developed by Lee (2018), one of the coauthors, and is specifically designed for text mining research in the field of economics and finance. For example, it can recognize words like '일드커브 (yield curve)' and '스티프닝 (steepening)' while the previous tools cannot. To encourage further research in this area and enhance comparability, eKoNLPy is made public at GitHub (https://github.com/entelecheia/eKoNLPy). Last but not least, in terms of future applications, our automated approach can be easily extended to measure other information such as macroeconomic uncertainty and stock market sentiments.

The rest of the paper is structured as follows. Section 2 reviews the related literature. Section 3 explains our data and methodology to develop text-based indicators that capture the sentiment of monetary policy. Section 4 empirically evaluates the performance of our text-based indicators and compare with other measures. Section 5 summarizes our finding and discusses future research venues.

2. Literature Review

In this section, rather than attempting an extensive literature review on text mining applied to economics, we limit our discussion relevant to central banking.[2] Earlier studies that use text mining approach rely more on the frequency of specific words, rather than on more sophisticated methods that measure tones or sentiments. For example, Choi and Varian (2012) show that the number of search keywords such as 'jobs' and 'welfare & unemployment' is a good predictor of leading indicators of US labor market as well as economic cycles. McLaren and Shanbhogue (2011) show that the volume of the related Google searches can predict changes in unemployment as well as house prices. A recent study by Baker et al. (2016) constructs economic policy uncertainty (EPU) measure by counting the number of news articles that contain specific words such as "uncertainty" and "economy" and shows that this measure is associated with reduced investment and employment in policy-sensitive sectors like defense, health care, finance, and infrastructure construction. At the macro level, innovations in policy uncertainty foreshadow declines in investment, output, and employment.

Concerning monetary policy, several studies attempt to extract additional information from social media or central bank communication. Using data from Twitter, Meinusch and Tillmann (2017) quantify beliefs about the timing of the exit from quantitative easing ("tapering") of market participants. Related to central bank communication, Lucca and

[2]See Gentzkow et al. (2017) for a survey of text mining approach to economic research and Bholat et al. (2015) for a survey more specific to central banking.

Trebbi (2011) build semantic scores using discussions of FOMC statements from news on days of announcements and find that longer-term Treasury yields mainly react to their semantic scores. Hansen and McMahon (2016) cluster sentences in FOMC communication with Latent Dirichlet analysis (LDA) and, within a factor augmented vector autoregression (FAVAR) framework, find that FOMC communication on forward guidance has stronger impact on the financial market. In line with our research to measure the sentiment of monetary policy, Picault and Renault (2017) manually classify all sentences in European Central Bank (ECB) press conferences and build a field-specific dictionary. Their measure of ECB monetary policy sentiment helps explain the current and future ECB monetary decisions and markets are more volatile when the Governing Council views on the economic outlook turned out to be negative. Nyman, Kapadia, Tuckett, Gregory, Ormerod, and Smith (2018) attempt to extract high-frequency sentiment from Bank of England's daily market commentary and show that the sentiment series track down the commonly used measure of volatility such as VIX very well and even moves as a leading indicator ahead of it.

Text mining is also applied to the area of financial stability. Based on over 1,000 releases of FSRs (Financial Stability Reports) and speeches/interviews by central bank governors from 37 central banks and over the past 14 year, Born, Ehrmann, and Frtzscher (2014) find that FSRs with optimistic tones lead to significant and lasting positive stock market returns, whereas there is no such effect for pessimistic ones. Nopp and Hanbury (2015) analyze CEO letters and annual management reports of 27 banks under ECB supervision. They find that sentiment scores of the textual data well reflect major economic events as well as the aggregate Tier 1 capital ratio evolution. Bholat, Brookes, Cai, Grundy, and Lund (2017) analyzes confidnetial letters sent by Prudential Regulation Authority (PRA) to banks and financial firm under supervision using a machine learning method. They find that, in terms of negative words and direct languages, the letters vary depending on the riskiness of the firm.

While we reckon that text mining approach that uses the Korean language in economic analysis is at the very early stage and there are few studies, we find two exceptions: Won, Son, Moon, and Lee (2017) and Pyo and Kim (2017). Won et al. (2017) provide a useful guidance on future research direction by comparing the methods of bag-of-words and word2vec and showing the outperformance of word2vec. However, as Won et al. (2017) point out, word2vec has the problem of often classifying antonyms as similar words. We address this problem by using n-gram embedding. A research of Pyo and Kim (2017) is notable because it is the first study to attempt the sentiment analysis to economic analysis using the Korean language. Pyo and Kim (2017) construct the sentiment index of financial markets using news articles and show that these investor sentiment measures are statistically associated with asset prices

such as government bond yields and exchange rate. While both Pyo and Kim (2017) and our study use the sentiment analysis, there are several points of departure. First, we use n-grams to capture the subtlety of texts by considering contexts. Second, we use a field-specific dictionary specifically designed for text mining in economics and finance. Third, we also construct the sentiment index based on market approach, in addition to lexical approach.[3]

3. Data and Methodology

According to Liu (2009), sentiment analysis is a series of methods, techniques, and tools about detecting and extracting subjective information, such as opinion and attitudes, from language. The roots of sentiment analysis are in the studies on public opinion analysis at the beginning of 20th century and in the text subjectivity analysis performed by the computational linguistics community in 1990's. The current form of sentiment analysis becomes popular with the advancement of computer technology and the availability of texts on the web.[4] While sentiment analysis has a long history in linguistics and political science, it has been utilized in economics rather recently. Especially, empirical studies in economics using the Korean language are rare.

Sentiment analysis generally takes the following processes: (i) preparing the corpus of interests, (ii) pre-processing texts, (iii) feature selection, (iv) polarity or sentiment classification of features and (v) measuring sentiments of sentences and documents. We briefly explain what we do in each step. Table 1 and Figure 1 summarizes our discussion in this section.

3.1. Preparing the Corpus

We collect 231,699 documents for the period of May 2005–December 2017, which include 151 minutes of MPC meetings, 206,223 news articles, and 26,284 bond analyst reports. Table 2 shows the types and numbers of documents and the average and maximum number of sentences. While our target texts are the MPC minutes, we use a large amount of other documents to build field-specific lexicons.

MPC Minutes The MPC minutes, recording discussions during MPC meetings, are released at 4 p.m. on the first Tuesday two weeks after each meeting since September 2012.[5]

[3]We explain our application of text mining approach in the following section, including the market and lexical approach.

[4]See Liu (2009) for a review on the evolution of sentiment analysis.

[5]Because of this convention of disclosing the minutes after two weeks after the market closes, it is difficult to perform event studies that attempt to guage the market impact of monetary policy. They were released

It consists of several sections:

- Outline provides the information about the administrative details
- Summary of discussion on the current economic situation contains the MPC members' discussion on economic situation, FX and international finance, financial markets, and monetary policy.
- Discussion concerning monetary policy decision records the views of individual members.
- Result of deliberation on monetary policy

We download the files of MPC minutes from May 2005 to December 2017 (151 minutes) from the BOK website.[6] We use only the second and third sections. Panel (a) and (b) in Figure 2 display the number of sentences in MPC minutes for each section over time. The length of the minutes has increased after the global financial crisis.

News Articles We collect news articles that include the word 'interest rates (금리)' from Naver and Infomax from January 2005 to December 2017.[7] They contain the information on the general economy, monetary policy, financial market, and public perception on the BOK's future monetary policy stances. We use only the articles from the top 3 news agencies (in terms of number of articles produced) for there are many duplicate articles from the originators. The number of news articles for our final use is 206,223. Among them, 42% (86,538) are from Yonhab Infomax, 33% (68,728) from EDAILY, and 25% (50,957) from Yonhab News. We remove the header and footer from the articles. Panel (c) and (d) in Figure 2 show the number of news articles over time.

Bond Analysts' Reports We also use bond analysts' reports for two reasons. One is that bond analyst reports show the experts' view on the monetary policy and the bond market. The other is to incorporate the informal styles of writing to our lexicons. Generally, bond analysts write in more informal ways than journalists do. We obtain the reports from WIEfn, a financial information service provider in Korea.[8] Panel (e) of figure 2 shows the number of reports from January 2005 to December 2017.

Our corpus is not only large in size but also covers various topics. Figure 3 shows the various topics of our corpus, which we extract using Latent Dirichlet Allocation (LDA) method, a topic modeling method. Table 3 shows the relative frequencies of the topics.

6 weeks after each meeting during April 2005 to September 2012.

[6] http://www.bok.or.kr/portal/singl/crncyPolicyDrcMtg/listYear.do?mtgSe=A&menuNo=200755

[7] https://news.naver.com, http://news.einfomax.co.kr

[8] https://www.wisereport.co.kr

3.2. Pre-processing Texts

3.2.1. Typical Steps of Pre-Processing

Pre-processing texts includes tokenization and normalization. Tokenization is a step to split longer strings of text into smaller pieces, or tokens, which are generally words. It can incorporate part of speech (POS) tagging, which assign a part of word such as nouns, verbs, adjectives, etc. Normalization is the process of transforming a text into a single canonical form. It includes the following: removing punctuation, stop words removal, converting numbers to their word equivalents, stemming, lemmatization, and case folding.[9]

A typical text pre-processing procedure for English are (i) converting all words to lower case, removing numbers and punctuation (ii) using a Porter (1980) stemming algorithm to reduce inflected words to their word roots (e.g, "increasing" to "increas", "unemployment" to "unemploy") or lemmatization (e.g, "better" to "good") , and (iii) removing stop words (e.g, a, the, an, of, to, etc.).

3.2.2. Korean NLP Python Library for Economic Analysis (eKoNLPy)

To convert Korean text into numerical expressions (e.g., bag of words, word embedding, etc.), there are several issues. The first issue is related to spacing. Unlike English, postpositions are not space-delimited and spacing rules are not strictly observed. Second, there are many foreign words that do not follow the foreign language notation standards. And many of them are field-specific. Third, there are various notations for the same-meaning words (e.g., inflation for '인플레이션', '인플레', and '물가'). This issue can be important when one uses n-grams. Various notations of synonyms increase the number of word combinations and diluting the frequency of n-grams. Fourth, lots of verbs and adjectives conjugate irregularly. Irregular conjugation also aggravates the explosion of dimension in n-gram models, which hinders polarity classification. The first issue of spacing is relatively well taken care of by currently available Korean morpheme analyzers. For example, one can use KoNLPy.[10] However, other issues are not. This is why we use eKoNLPy by Lee (2018), developed by one of the coauthor. eKoNLPy constructs a dictionary specific to economics and finance and uses its own morphologocial analyzer.

Related to the second issue, to fully support economics and finance domain-specific terms

[9]Stop words removal is to drop stop words such as 'it', 'the', 'etc' and others. Stemming is to count just stems (for example, using 'bank' for 'banking' and 'banks'). Lemmatization is to group the inflected forms of words so that they can be analyzed as a single item. POS tagging often helps lemmatization. For example, 'saw' can be the past tense of a verb 'see' or a noun.

[10]KoNLPy is a Python package for natural language processing (NLP) of the Korean language (http://konlpy.org/en/v0.5.1/references).

(i.e., jargon and foreign words), eKoNLPy is equipped with pre-supplied 4,202 field-specific terms that are acquired from readily available economic term dictionary on the internet.[11] And it has the functionality to easily add custom terms and foreign words to the dictionary for POS tagging. For the third issue, to deal with various notations of synonyms, eKoNLPy has pre-defined 1,325 pair of synonyms in the dictionary and supports the function of replacing synonyms. The last issue, conjugation of adjectives and verbs, can be handled by stemming or lemmatization. Normalizing irregular conjugation of Korean words can be better addressed by lemmatization, rather than by stemming, since lemmatization consider the morphological analysis of words. To deal with this problem, eKoNLPy supports lemmatization of 1,291 adjectives and verbs, which are frequently used in economic and finance domain.

Since eKoNLPy is developed for the purpose of text mining for economic analysis from the beginning, we expect its outperformance in economic analysis, compared to KoNLPy. For example, consider the following sentence:

> "한국은행이 12일 금융통화위원회(금통위) 회의를 열고 기준금리를 현행 연 1.50 로 동결했다."

We find that eKoNLPy successfully recognizes the words of '금융통화위원회 (Monetary Policy Committee)' and '금통위 (MPC)' while KoNLPy, which uses a general-purpose dictionary, cannot.[12] Consider the following phrase on bond market:

> "금리 박스권 상단 상향과 일드 커브 완만한 스티프닝 전망"

eKoNLPy recognizes '일드커브 (yield curve)' and '스티프닝 (steepening)' while KoNLPy cannot.

3.3. Feature Selection

Since not all words are used to express opinions, feature selection is necessary to restrict words or phrases to a targeted list of words that express opinions. Restricting words also facilitate the speed of processing by reducing the dimension of term (word) vectors. Meanwhile,

[11] There are several online economic term dictionaries available at Naver, Maekyung, Hankyung, etc.

[12] The result from KoNLPy is the following:
'한국은행/NNP', '이/JKS', '12/SN', '일/NNBC', '금융/NNG', '통화/NNG', '위원회/NNG', '(/SSO', '금/NNG', '통/NNG', '위/NNG', ')/SSC', '회의/NNG', '를/JKO', '열/VV', '고/EC', '기준/NNG', '금리/NNG', '를/JKO', '현행/NNG', '연/NNG', '1/SN', './SY', '50/SN', '%/SY', '로/JKB', '동결/NNG', '했/XSV', '다/EF', './SF'
The result from eKoNLPy is as follows:
'한국은행/NNG', '이/JKS', '공공요금/NNG', '12/SN', '일/NNG', '금융통화위원회/NNG', '금통위/NNG', '회의/NNG', '를/JKO', '열/VV', '고/EC', '기준금리/NNG', '를/JKO', '현행/NNG', '연/NNG', '1/SN', './SY', '50/SN', '%/SY', '로/JKB', '동결/NNG', '했/XSV', '다/EC'

single words often lose the context. For example, while the word "recovery" in isolation appears to carry a positive message, the phrase "sluggish recovery" does not.[13] When positive and negative words are combined, like a bi-gram phrase "lower unemployment," the sentiment is not easy to measure. To address this problem, we decide to use n-grams.[14] However, increasing the length of n-grams has a trade-off. With too long n-grams (say, 10-grams), we might fall into the problem of over-fitting to the sample; as the lexicons are too much specific to the target documents, it is hard to apply that lexicons to other types of documents such as news articles or experts' writings. In addition, a curse of dimensionality arises with n-grams.[15] There is an exponential growth in the number of features, and the probability of seeing the n-grams with the same features gets much smaller. This explosion of dimension also causes computational problems regrading memory size and speed of processing.

To address this trade-off, we set the n of n-gram to 5 with additional rules. To avoid the explosion of dimension, we use the limited word set in forming n-grams by limiting the part-of-speech tag of words to nouns (NNG), adjectives (VA, VAX), adverbs (MAG), verbs (VA), and negations.[16] We also drop n-grams that occur less frequently than 15 times.[17]

The final word set is comprised of 2,712 words and we obtain the resulting 73,428 n-grams. Note that, since we cover from 1-gram to 5-gram, our n-grams naturally include single words (1-grams).[18] The next step is to classify polarity of these n-grams to use them for measuring sentiments of sentences or documents.

[13] Apel and Grimaldi (2014) use two-word combinations (bi-grams) of a noun and an adjective, such as "higher inflation" or "slower growth," to make the hawkish-dovish classification.

[14] Picault and Renault (2017) define the field-specific lexicon by considering n-grams (from 1-gram to 10-grams) appearing at least twice in their sample. They classify the polarity of n-grams by calculating the probability that they belong to which category of sentences (dovish, neutral, hawkish or positive, negative, neutral), after classifying manually all sentences pronounced during ECB introductory statements. By limiting n-grams to those having a probability over 0.5 in each class, their final field-specific lexicon is composed of 34,052 n-grams.

[15] As text is represented as very high-dimensional but sparse vectors, it is challenging to reduce dimensionality while preserving the important variation across documents. With the introduction of n-gram, which is a contiguous sequence of n words in a text, this problem is aggravated. With a 1,000 unique word corpus, a bigram model needs $1,000^2$ values; a trigram model will need $1,000^3$; and a 4-gram will need $1,000^4$.

[16] Appendix B shows the eKoNLPy tagest for POS tagging.

[17] From the perspective of n-grams being effective opinion-bearing features for sentiment analysis, ideal n-grams should contain following factors: the agent or doer (i.e., the person or entity who maintains or causes that affective state), the target (i.e., the entity about which the affect is felt), how (i.e., the direction or the degree of affective state), and negation if any.

[18] We also check the sensitivity of n in n-gram. we find that a higher n increases in-sample performance and lowers out-of-sample performance in terms of the accuracy of polarity classification, which we discuss in section 3.4. As we explain above, this result suggests that a higher n makes n-gram more document-specific.

3.4. Polarity Classification

If there are no well-known lists of polarity words like Harvard-IV or LM dictionary, we need to classify the polarity of our selected features (n-grams in our case) on our own.[19] There can be several categories of polarity classification. First, there are supervised vs. unsupervised (automated) approaches depending on whether it needs human intervention or not. Google Cloud Sentiment Analysis API is an example of supervised classification in which the classifier is trained on the massive amount of documents. An example of unsupervised approach is semantic orientation using PMI (Pointwise Mutual Information) to measure the similarity between words and polarity proto types.[20]

Second, there are machine-learning-based vs. lexical-based methods depending on whether it employs opinion words for classification. In the lexical-based approach, there are three methods to obtain the polarity word list: manual, dictionary-based, and corpus-based. The manual method is very time-consuming and prone to human errors.[21] The dictionary-based approach searches the dictionary starting from the seed words to find their synonyms and antonyms. This approach requires a well constructed lexical database such as WordNet or thesaurus.[22] The disadvantage of this approach is its inability to find field-specific polarity words. The corpus-based method finds polarity words by searching patterns that occur together along with seed words in a large corpus. This approach has a major advantage in that it can find field- and context-specific sentiment words and their polarities using a corpus of that domain. This fact makes the corpus-based approach the most suitable for economics or finance area where jargon or words with different connotations are prevalent.

We classify polarity of n-grams in two ways. One is the market approach that classifies polarity from market information using machine learning. The other is the corpus-based approach that classifies polarity using word (n-gram in our case) embedding and seed words,

[19]For unigrams (single words), the oldest is the General Inquirer (Stone, Dunphy, and Smith, 1966) also known as Harvard IV-4, which has many categories of word lists including 1,915 words of positive outlook and 2,291 words of negative outlook. In financial context, negative words are mainly used for sentiment analysis (Tetlock, 2007). A widely used word list in finance literature is that of Loughran and McDonald (2011), which has lists of single words by category (Negative, Positive, Uncertainty, Litigious, Modal, Constraining). Their research indicates that the LM dictionary has a better correlation with financial metrics. The LM dictionary is available at the following link: https://sraf.nd.edu/textual-analysis/resources/

[20]Semantic orientation is a concept from computational linguistics and defines the position of a word or string of words between two opposite concepts. Lucca and Trebbi (2011) and Tobback, Nardelli, and Martens (2016) employ this SO-PMI (Semantic Orientation from Point-wise Mutual. Information) method to measure the sentiments of the FOMC statements and the ECB's press statements. They measure their sentiment indicators by counting the co-occurrences of words or strings in their documents with words that are normally associated with the sentiments of monetary policy (dovish or hawkish) using Google Search.

[21]Hence, it is used in the specific tasks with only small focused list of words like the cases of Apel and Grimaldi (2014) and Bennani and Neuenkirch (2016).

[22]WordNet® is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. https://wordnet.princeton.edu

which we call lexical approach.[23] We will explain these two approaches one by one.

3.4.1. Market Approach

There are many attempts to extract information about market expectations from asset prices, as the survey of Söderlind and Svensson (1997) shows. To the extent that financial market is efficient and asset prices reflect information of financial market, one can extract information. Then, why can't one extract information from text? With this in mind, we classify the polarity of features using market information and call it "market approach."

To determine the relative weights of words, Jegadeesh and Wu (2013) use words as the dependent variable and stock returns as the explanatory variable. If the coefficient of a word is positive and large, than the word has higher weight. That is, they estimate market reactions and use it as relative weights. One advantage of Jegadeesh and Wu (2013) is that it does not rely on subjective judgment. The difference with our market approach is that they start from the known word lists like Harvard IV-4 or LM dictionary and use regression-based approach to determine the relative weights (the sign and the magnitude) of the terms. And they use uni-gram, not n-grams. Our market approach also uses market information to determine the polarity of features. However, to extract n-grams from a large corpus and classify their polarity, we use the machine learning method.

For our market approach, we use the Naïve Bayes classifier (NBC), a simple probabilistic classifier. NBC is very simple but still competitive with more advanced methods including support vector machines. NBC is called naive, because it assumes that all features are independent given the class (hawkish or dovish in our case). Even though NBC is not a feature selection tool, this independence assumption make us use the conditional probability of each feature as a polarity score.

Machine learning methods, including NBC, are mostly supervised ones, which depend on the existence of labeled training documents. Training documents are usually obtained from the review by the public when it is available. Otherwise, some experts need to manually label training documents. The former is not available for monetary policy. The latter is labor- and cost-intensive, and is subject to experts' judgements. To circumvent these problems and exploit the information of financial market, we label news articles and reports in our corpus as hawkish (dovish) if the 1-month change of Call rates is positive (negative) on the day

[23] Word embedding is the collective name for a set of language modeling and feature learning techniques in natural language processing (NLP) where words or phrases from the vocabulary are mapped to vectors of real numbers. Conceptually it involves a mathematical embedding from a space with one dimension per word to a continuous vector space with a much lower dimension (https://en.wikipedia.org/wiki/Word_embedding).

they are released.[24]

We randomly divide our labeled sentences (more than 4 million sentences) into a training set and a test set by 9:1 ratio.[25] Using 5-grams (from 1-gram to 5-grams) as features for each sentence, we train the classifier and check its accuracy.[26] The trained NBC yields the conditional probability of each feature given the class (hawkish/dovish), which we use as a polarity score of the feature:

$$polarity\ score = \frac{p(feature|hawkish)}{p(feature|dovish)} = \frac{p(feature\&hawkish)/p(hawkish)}{p(feature\&dovish)/p(dovish)} \tag{1}$$

Roughly speaking, it is like labeling an *n-gram* as 'hawkish' if it presents itself more often in 'hawkish' *documents*, compared to 'dovish' ones.[27]

Because we use random sampling and a probabilistic classifier, every training yields different posterior probability of each class. To obtain better predictive performance, we repeat this procedure 30 times and use the average of the polarity scores as a final one. It is called bagging in machine learning.[28] While this is not the direct performance measure of our lexicon, the average accuracy of NBC is 86% (positive precision: 90%, positive recall: 84%, negative precision: 82%, negative recall: 88%).

We classify the polarity of our lexicon as hawkish (dovish) if the polarity score is greater (less) than 1, excluding lexicon in the grey area using intensity of 1.3 as a threshold.[29] The final number of lexicon is 18,685 for hawkish and 21,280 for dovish. A sample of polarity lexicon is provided in table 4.

3.4.2. Lexical Approach

While our market approach that associates the release dates of documents with changes in Call rates is not only easy but also effective to take advantage of market information, one

[24]The actual threshold is $\pm 3bp$ to exclude meaningless movements.

[25]One may suggest using three kinds of sets: training, test, and validation. Since we evaluate our polarity classification with out-of-sample documents, we do not need a validation set and our out-of-sample evaluation is more rigorous.

[26]In order to improve the accuracy of the classification and to avoid multiple counting, we consider only the highest n-gram when multiple overlapping n-grams are found in each sentence.

[27]Alternatively, one may consider the following:

$$polarity\ score = \frac{p(hawkish|feature)}{p(dovish|feature)} = \frac{p(feature\&hawkish)}{p(feature\&dovish)}$$

In our case, since $p(hawkish) = 0.53$ and $p(dovish) = 0.47$, these two formulas produce the similar result.

[28]Bootstrap aggregating, often abbreviated as bagging, involves having each model in the ensemble vote with equal weight.

[29]Intensity measures the relative strength of the polarity, which is a ratio of the conditional probability of the feature given the class (hawkish/dovish) with the greater of the two in numerator.

may suggest many different ways of incorporating market information. For example, one can use the stock market index instead of Call rates. Or one can measure the change of Call rates over one week, not one month. Rather than addressing all these possibilities, we construct another indicator that takes an opposite stance in that it does not use market information at all. Instead, it uses the seed words, which we call lexical approach.

Lexical approach is based on an intuitive observation: if two words appear together frequently in the same context, they are likely to have the same polarity. Then the polarity of an unknown word can be determined by calculating the relative frequency of co-occurrence with another word. This could be done by using the concept of PMI (Pointwise Mutual Information). One can use SO-PMI (Semantic Orientation from PMI) proposed by Turney (2002) for polarity classification.[30]

While this approach is quite intuitive, there are two problems. First, it sometimes fails to recognize antonyms because it judges the polarity based on co-occurrence. Second, the outcome is affected by choices of seed words. To address the first problem, we use ngram2vec by Zhao, Liu, Li, Li, and Du (2017) instead of word embedding. They show that n-gram embedding is effective in finding antonyms. For the second problem, We adopt the SentProp framework by Hamilton et al. (2016), a state-of-the-art domain-specific sentiment induction algorithm. The SentProp framework addresses this issue by bootstrapping seed words.

We place the seed set of words (n-grams in our case) and our n-grams in a vector space (lexical graph) and measure the proximity of our n-grams to this seed. The polarity of an n-gram is proportional to the probability of a random walk from the seed set hitting that n-gram. Each feature will have two probabilities; one for hawkish, the other for dovish. A final polarity score is the relative ratio of the two as in equation (1).

We train ngram2vec using the entire 232,658 documents of our corpus. The parameters we use for training are 5-gram for center words, 5-gram for context words, window size of 5, negative sampling size of 5, and 300 dimension for vector representation. Our corpus has 344,293 unique n-grams with a minimum frequency limit of 25, which yield 410,902,512 pairs of n-grams (21.7 GB in size). With this resulting n-gram vector, we bootstrap by running our propagation 50 times over 10 random equally-sized subsets of the hawkish and dovish seed sets. Table 5 shows seed sets.

Likewise the market approach, we classify polarity of our lexicon as hawkish (dovish) if

[30]PMI is a technique for quantifying the similarity between two random variables based on probability theory. Using PMI, the similarity between two lexicons is measured as:

$$PMI(w_1, w_2) = \log \frac{P((w_1, w_2)}{P(w_1)P(w_2)},$$

where w_1 and w_2 are the two lexicons under consideration.

the polarity score is greater (less) than 1, excluding lexicon in the grey area using intensity of 1.1 as a threshold. The final number of lexicon is 11,710 for hawkish and 12,246 for dovish. A sample of polarity lexicon is provided in table 6.

To see if the market and lexical approach give the similar result, we count the number of common n-grams. Given 39,965 n-grams from the market approach and 23,956 n-grams from the lexical approach, there are 14,154 common n-grams. Among them, 9,791 n-grams (69% of common n-grams) have the same polarity.

3.4.3. Evaluation

We check the accuracy of our lexicon classification. In principle, the criteria of judging the accuracy is how well the classification of sentiment agrees with human judgments.[31] We evaluate the accuracy using documents that are not used in building our lexicons.

Documents for evaluation are introductory statements from the BOK Governor's news conference about monetary policy decisions. With the documents from May 2009 to January 2018, we manually label 2,341 sentences as hawkish, neutral, and dovish. To check the consistency of our classification, we train a Naïve Bayes classifier with randomly selected 60% of hawkish and dovish sentences and test with the remaining sentences. With 30 times of iteration, the average accuracy of classifiers is about 86%, which we think is above par accuracy.

Finally, we check the accuracy of our lexicons using labeled sentences that is completely out-of-sample. For the lexicon generated by market approach, the accuracy is 68% (positive precision: 63%, positive recall: 75%, negative precision: 74%, negative recall: 62%). For the lexicons generated by lexical approach, the accuracy is 67% (positive precision: 69%, positive recall: 71%, negative precision: 65%, negative recall: 62%).

3.5. Measuring Sentiments

With the lexicons in hand, the last step is to measure the tone of our target documents. We adopt a two-step approach to measure the tone of documents. First, we calculate the tone of a sentence based on the number of hawkish and dovish features (n-grams) in each sentence. Specifically, the tone of a sentence s is defined by following formula:

$$tone_s = \frac{No.\ of\ hawkish\ features - No.\ of\ dovish\ features}{No.\ of\ hawkish\ features + No.\ of\ dovish\ features} \tag{2}$$

[31] According to research by Amazon's Mechanical Turk, human raters typically only agree about 79% of the time (https://mashable.com/2010/04/19/sentiment-analysis/#skdJb2rbx5qg).

Then, we calculate the tone of a document i by following formula:

$$tone_i = \frac{No.\ of\ hawkish\ tone_{s,i} - No.\ of\ dovish\ tone_{s,i}}{No.\ of\ hawkish\ tone_{s,i} + No.\ of\ dovish\ tone_{s,i}} \tag{3}$$

It creates a continuous variable $tone_i$ for each document, which is bound between -1 (dovish) and $+1$ (hawkish).[32] This index is our indicator of quantifying the sentiment of monetary policy. We denote the indicator based on market approach and lexical approach by $tone^{mkt}$ and $tone^{lex}$, respectively. We examine their statistical properties and explanatory power of the current and future monetary policy decisions in the following section.

4. Empirical Analysis

In this section, we attempt to answer the following questions:

1. Can our lexicon-based indicators ($tone^{mkt}$ and $tone^{lex}$) explain the BOK's current and future monetary policy decisions? In particular, do they have additional information that are not available in the existing macroeconomic data?

2. Is it important to use a field-specific dictionary?

3. Is it important to use the original Korean text, not Korean-to-English text?

For the first question, our answer is astoundingly "yes." We consider an augmented Taylor rule to compare the explanatory powers of our indicators, macroeconomic variables, and other proxies related to economic uncertainty. We find that our indicators have additional explanatory power in addition to macroeconomic variables. For the second and third question, we also show that it is crucial to stick to the original Korean text and use a dictionary specific to economics and finance terminologies.

4.1. Measures of MP Sentiment

Based on the methodology we discuss in section 3, we develop lexicon-based indicators that capture the sentiment (or tone) of the BOK MPC's minutes: $tone^{mkt}$ and $tone^{lex}$. The former uses the market approach while the latter uses lexical approach.

[32] This way of measuring tones is similar to that of diffusion index in that it only takes account of direction. Kennedy (1994) find that the diffusion indexes for industrial production and employment have predictive power in a twenty-five year sample beginning in 1967. From the perspective of signal-extraction, by ignoring the magnitudes, this way of calculation insulates it from idiosyncratic shocks and offers a cleaner view of the persistent component.

Figure 4 shows the time-series of $tone_t^{mkt}$ and $tone_t^{lex}$ with those of the BOK policy rate, other measures of economic uncertainty. Panel (a) in figure 4 shows the time-series of $tone_t^{mkt}$ and $tone_t^{lex}$. They move closely with each other. The correlation coefficient between the two indicators are 0.85. This co-movement is interesting given that these two indicators are constructed very differently. Panel (b) shows that our indicator $tone_t^{mkt}$ well captures the movements of the BOK policy rate. Panel (c) and (d) compare our indicator with other measures such as economic policy uncertainty index by Baker et al. (2016), and uncertainty index by Jurado et al. (2015).[33] Panel (c) shows that, except the period of the recent financial crisis, our indicator and economic policy uncertainty index do not seem to move in opposite direction. The correlation coefficient between the two is only -0.06. In contrast, panel (d) shows that our indicator is negatively associated with uncertainty index. The correlation coefficient is -0.54.

Figure 5 compares our indicator $tone^{mkt}$ with monetary policy decisions (MPD_t), output gap ($y_t - y_t^*$), CPI (π_t), and stock market index. Following Picault and Renault (2017), to account for the BOK's non-standard monetary policy, MPD_t takes a value of zero when there is no change in monetary policy stance, $+1$ for a hawkish monetary policy decision (an increase of policy rate by 25 basis points), -1 for a dovish monetary policy decision either through an increase of policy rate or a non-standard policy, and -2 for a very dovish decision with both a policy rate cut and a non-standard policy.[34] Output gap ($y_t - y^*$) is defined as the difference between the industrial production and its trend from Hodrick-Prescott filter. Panel (a) clearly shows that our indicator well tracks the changes in monetary policy stance that take into account BOK's non-standard policy. Panel (b)–(d) shows the time-series of output gap, CPI, and stock market index (KOSPI). Correlation coefficient of $tone^{mkt}$ with monetary policy decision, output gap, CPI, and KOSPI are 0.58, 0.40, 0.40, and -0.36, respectively.

Panel (a) in Figure 6 shows the correlation coefficients among the selected variables. It clearly shows that the BOK policy rate and industrial production have a strong negative correlation. BOK policy rate has a positive correlation with inflation (π) and $tone^{mkt}$, while it has a negative correlation with economic policy uncertainty measure of US and South Korea. In the following section, we formally test the explanatory power of various indicators in an augmented Taylor rule specification.

[33] For both measures of economic policy uncertainty and uncertainty, we use the Korean versions. As of August 2018, one can use economic policy uncertainty indexes of 25 countries at http://www.policyuncertainty.com. For the Korean version of Jurado et al. (2015), refer to Shin, Zhang, Zhong, and Lee (2018) and the journal website.

[34] Appendix A shows the chronological list of the BOK's non-stanard monetary policy. Source: Bank of Korea website (www.bok.or.kr).

4.2. Explaining the BOK's Moneary Policy Decisions

In order to assess the relation between the BOK MPC's policy rate decisions and the information content of MPC minutes, we test the explanatory power of our lexicon-based indicators ($tone_t^{mkt}$ and $tone_t^{lex}$) on both contemporaneous and future decisions. We consider the same specification used in Picault and Renault (2017). They use an ordered probit model to estimate the coefficients of the forward-looking Taylor rule and compare the explanatory powers of their own lexicon-based indicators with those of macroeconomic variables and other types of indicators. They start from the following form:

$$MP_t = \alpha + \rho MP_{t-1} + \gamma_1(\pi_t - \pi^*) + \gamma_2(y_t - y_t^*) + \gamma_3\pi_t^e + \gamma_4 y_t^e + \epsilon_t, \tag{4}$$

where MP_t is the BOK's monetary policy stance and $(\pi_t - \pi^*)$ is the inflation gap defined as the difference between CPI and the inflation target $\pi^* = 2\%$. $(y_t - y^*)$ is the output gap. π_t^e is the expected inflation of Consumer Survey Index and y_t^e is the leading Economic Composite Index.[35] We include the lagged variable MP_{t-1} to control for the smoothing of monetary policy.

For estimation, in order to ensure stationarity, we estimate the differenced version of equation (4) with adding a variable X_t[36]:

$$\begin{aligned}\Delta MP_t = \rho\Delta MP_{t-1} + \gamma_1\Delta(\pi_t - \pi^*) + \gamma_2\Delta(y_t - y_t^*) \\ + \gamma_3\Delta\pi_t^e + \gamma_4\Delta y_t^e + \beta X_t + u_t,\end{aligned} \tag{5}$$

where X_t is the variable that potentially help explain changes in monetary policy stance in addition to macroeconomic variables. For X_t, we consider our lexicon-based indicators ($tone_t^{mkt}$ and $tone_t^{lex}$), economic policy uncertainty index (EPU_t) by Baker et al. (2016), and uncertainty index (UI_t) by Jurado et al. (2015). For changes in monetary policy stance (ΔMP_t), we use two variables following Picault and Renault (2017). First, we focus on changes in policy rate (ΔBOK_t). During the sample period, the BOK MPC increases its policy rate by 25 basis points on thirteen occasions and decreases it on thirteen occasions (nine times by 25 basis points, twice by 50 basis points, and twice by 100 basis points).[37]

[35] Statistics Korea (http://kostat.go.kr) reports three kinds of Economic Composite Index: leading, coincident, and lagging. The leading index is based on nine indicators that are closely related to future economic activities: construction orders received, opening-to-application ratio, inventory circulation indicator, consumer expectation index, machinery orders received, import of capital goods, stock price index, total liquidity, interest rate spread, 3-year treasury bonds less call rate, and net barter terms of trade.

[36] Using the augmented Dickey-Fuller test, we find that policy rate (BOK_t), inflation gap ($\pi_t - \pi^*$), output gap ($y_t - y_t^*$), expected inflation (π_t^e), output leading indicator (y_t^e), uncertainty index (UI_t) have a unit root while $tone_t^{mkt}$, $tone_t^{lex}$ and EPU_t turn out to be stationary.

[37] Since we use the monthly data for our empirical analysis, we measure the changes in policy rate on a

Thus ΔBOK_t takes values of -1.0, -0.5, -0.25, 0, and $+0.25$ in our sample. The other is a variable of monetary policy decision (MPD_t) to account for the BOK's non-standard monetary policy. In a forward-looking approach, equation (5) is rewritten as:

$$\begin{aligned}\Delta MP_{t+k} = \rho \Delta MP_t + \gamma_1 \Delta(\pi_t - \pi^*) + \gamma_2 \Delta(y_t - y_t^*) \\ + \gamma_3 \Delta \pi_t^e + \gamma_4 \Delta y_t^e + \beta X_t + u_t. \end{aligned} \tag{6}$$

We use $k = 1$ and $k = 2$ and report the case of $k = 2$ for brevity.

If the MPC minutes do provide any information additional to previously released macroeconomic data, then our indicators of $tone^{mkt}$ and $tone^{lex}$ should be significant in equation (5) and (6). In addition, we expect a positive coefficient: more hawkish (dovish) sentiment should be associated with tight (loose) monetary policy. Table 7 shows the estimation result when we measure the change in monetary policy stance (ΔMP_t) as the change in policy rate (ΔBOK_t). It shows that our indicators $tone^{mkt}$ and $tone^{lex}$ are highly significant, while measures of overall economic uncertainty (EPU_t and UI_t) are not. Moreover, adding $tone^{mkt}$ or $tone^{lex}$ noticeably raises the value of R^2. Comparing column (2) and (3), adding $tone^{mkt}$ raises R^2 from 0.095 to 0.446. When the dependent variable is ΔBOK_{t+2}, we observe the similar result. The result from table 7 strongly suggests that our lexicon-based indicators contain additional information that are not captured by macroeconomic data.[38]

Table 8 shows the result when the dependent variable is MPD_t that considers BOK's non-standard policy. We obtain the similar results. Both $tone^{mkt}$ and $tone^{lex}$ are highly significant with considerably raising R^2, while other measures are not.[39]

To check the robustness of our result, we also estimate the different specification. Apel and

monthly basis, not on each meeting. For example, a regular MPC meeting on October 9, 2008 lowered the rate by 25 basis points and an emergency MPC meeting on October 27, 2008 lowered the rate by 75 basis points. We record $\Delta BOK_t = -100$ basis points in October, 2008.

[38] We also run a horse race with adding all three variables of $tone^{mkt}$, EPU_t and UI_t. The estimation result is as follows (the numbers in parentheses are standard errors):

$$\begin{aligned}\Delta \hat{MP}_{t+2} = \underset{(0.88)}{-0.06}\Delta MP_t + \underset{(0.39)}{0.20}\,\Delta(\pi_t - \pi^*) + \underset{(5.24)}{5.15}\,\Delta(y_t - y_t^*) + \underset{(0.97)}{0.85}\,\Delta \pi_t^e \\ + \underset{(0.49)}{0.11}\,\Delta y_t^e + \underset{(0.41)}{1.58}\,tone_t^{mkt} + \underset{(0.003)}{0.001}\,EPU_t \underset{(2.36)}{-4.70}\,UI_t, \qquad \text{pseudo } R^2 = 0.22.\end{aligned}$$

Note that $tone^{mkt}$ is still highly significant and R^2 does not increase much with adding EPU_t and UI_t. When we use $tone^{lex}$ instead of $tone^{mkt}$, we obtain the similar result:

$$\begin{aligned}\Delta \hat{MP}_{t+2} = \underset{(0.84)}{-0.59}\Delta MP_t + \underset{(0.38)}{0.16}\,\Delta(\pi_t - \pi^*) + \underset{(5.18)}{6.47}\,\Delta(y_t - y_t^*) + \underset{(0.96)}{0.58}\,\Delta \pi_t^e \\ + \underset{(0.47)}{0.01}\,\Delta y_t^e + \underset{(0.57)}{1.92}\,tone_t^{lex} + \underset{(0.00)}{0.003}\,EPU_t \underset{(2.23)}{-0.95}\,UI_t, \qquad \text{pseudo } R^2 = 0.19.\end{aligned}$$

[39] We also run the same specification by OLS and obtain the similar result.

Grimaldi (2014) use the following specification and estimate the coefficients using ordered probit:

$$\Delta r_{t+1} = \alpha_1 \Delta r_t + \alpha_2 X_t + \alpha_3 GDPgrowth_t + \alpha_4 Inflation_t + \epsilon_{t+1}.$$

If macroeconomic variables contain all the relevant information for the next policy rate, then the estimate of α_2 would be largely insignificant. We estimate the same regression equation with or without $tone_t^{mkt}$ after replacing GDP growth rate with IP growth rate for monthly estimation. We obtain the following result. The numbers in parentheses are standard errors.

$$\hat{\Delta r_{t+1}} = \underset{(0.65)}{1.90}\Delta r_t + \underset{(4.64)}{7.28} IPgrowth_t + \underset{(0.10)}{0.12} CPI_t, \quad \text{pseudo } R^2 = 0.08.$$

With $tone_t^{mkt}$, we obtain

$$\hat{\Delta r_{t+1}} = \underset{(0.90)}{-1.67}\Delta r_t + \underset{(0.86)}{4.20} tone_t^{mkt} + \underset{(5.33)}{9.73} IPgrowth_t \underset{(0.14)}{-0.28}\, CPI_t, \quad \text{pseudo } R^2 = 0.37.$$

Again, our indicator $tone_t^{mkt}$ is highly significant and raises $R^2 = 0.37$ from $R^2 = 0.08$. This result strongly suggests that our indicator $tone_t^{mkt}$ contain the relevant information on the future monetary policy stance beyond information in macroeconomic variables.

4.3. Comparison with Other Text-Based Indicators

One may wonder if the original Korean text should be used. What if one translates Korean text into English and applies the standard procedures for English text. It is a legitimate question given the availability of more advanced and diverse text mining techniques for English text. Further, is it really important to use a field-specific dictionary? In order to answer these questions, we compare our indicators with four other indicators: an indicator that specializes in Korean texts and uses a general-purpose dictionary ($tone^{ksa}$), and three English-based indicators ($tone^{google}$, $tone^{HIV4}$, and $tone^{LM}$). $tone^{ksa}$ is based on KKMA (Kind Korean Morpheme Analyzer) project developed by SNU (Seoul National University) Intelligent Data Systems Laboratory, which is one of the most popular tools for analyzing Korean texts.[40] But it uses a general-purpose dictionary, not like $tone^{mkt}$ and $tone^{lex}$ that uses a dictionary specific to the field of economics and finance. For English-based text analysis, we translate all the MPC's minutes into English using Google Cloud Translation.[41] $tone^{google}$ measures the tone of minutes using the service of sentiment analysis provided by Google

[40] See Lee, Yeon, Hwang, and Lee (2010) for more detail.

[41] https://cloud.google.com/translate/.

Cloud Natural Language.[42] $tone^{HIV4}$ is based on the general-purpose Harvard IV-4 dictionary and $tone^{LM}$ is based on the field-specific dictionary of Loughran and McDonald (2011).

Figure 7 shows the time-series of $tone^{ksa}$, $tone^{google}$, $tone^{HIV4}$, and $tone^{LM}$ with that of $tone^{mkt}$. Panel (a) shows that $tone^{ksa}$ has much less variations compared to $tone^{mkt}$. It does not fluctuate much even during the period of the recent financial crisis. While English-based indicators have more variation compared to $tone^{ksa}$, it seems that the degree of comovement with $tone^{mkt}$ is rather weak. Panel (b) in Table 6 shows the correlation coefficients. The correlation coefficients of $tone^{mkt}$ with $tone^{ksa}$, $tone^{google}$, $tone^{HIV4}$, and $tone^{LM}$ are 0.08, 0.26, 0.10, and 0.39, respectively. While both $tone^{mkt}$ with $tone^{ksa}$ are based on the original Korean texts, the statistical association is very low.

We compare the performance of these indicators based on equation (5) and (6) when performance is measured by the statistical significance of an individual coefficient and R^2. Our conjecture is:

(i) $tone^{mkt}$ outperforms $tone^{ksa}$ in terms of explaining changes in the current and future BOK policy rate.

(ii) $tone^{LM}$ outperforms $tone^{google}$ and $tone^{HIV4}$.

(iii) $tone^{mkt}$ outperforms $tone^{LM}$, $tone^{google}$, and $tone^{HIV4}$.

Roughly speaking, (i) and (ii) are to test the benefit of using a field-specific dictionary and (iii) is to compare the results from Korean texts and Korean-to-English texts.

Table 9 shows the estimation result. When the dependent variable is the current change in BOK policy rate (ΔBOK_t), column (1) and (2) show that, $tone^{ksa}$ fails to capture the change in BOK policy rate, while it is based on the most popular tool for the Korean text analysis. Column (3)–(5) shows that, while $tone^{google}$ and $tone^{HIV4}$ are not statistically significant, $tone^{LM}$ is statistically significant at 10% level. These results suggest that it is important to use a field-specific dictionary. When we compare $tone^{mkt}$ (column (1)) with $tone^{LM}$ (column (5)), R^2 in column (1) is far higher than R^2 in column (5), attesting the importance of the original Korean text. When the dependent variable is the future change in BOK policy rate ΔBOK_{t+2}, $tone^{mkt}$ outperforms other indicators in terms of its statistical significance and R^2. Considering our discussion based on figure 7 and table 9, it is desirable to use the original Korean text with a field-specific dictionary. And it also confirms the validity of our approach to quantify the information in the MPC minutes.

[42] https://cloud.google.com/natural-language/. Both Google Cloud Translation and Google Cloud Natural Language are the part of the Google's AI and Machine Learning products.

5. Concluding Remarks

We develop text-based indicators that quantify the sentiment of monetary policy using a field-specific Korean dictionary and n-grams. We show that our indicators help explain the current and future monetary policy decisions and perform better compared to other indicators. We also show that using a field-specific dictionary and the original Korean text is important, which would benefit future research in this field.

Our empirical results suggest future research venues. First, it is important to examine what kinds of information our indicators do (or do not) have compared to the BOK policy rate and macroeconomic variables. If we interpret the BOK policy rate as a threshold variable or a latent variable of our indicator of monetary policy sentiment, it would be interesting to feed our measure into the standard VAR systems or DSGE models that analyze the effect of monetary policy. Alternatively, it would be interesting to examine the implication of the discrepancy between our indicator and policy rate. As shown in Panel (b) of figure 4, our indicator $tone^{mkt}$ and the BOK policy rate started diverging from 2016. While $tone^{mkt}$ becomes more hawkish, the policy rate does not. One can examine the information content in $tone^{mkt}$ orthogonal to the policy rate. Another direction would be Hansen and McMahon (2016). They construct two separate indicators on the state of economy and forward guidance from the Fed statements and use a Factor-Augmented VAR (FAVAR) to examine how these two dimensions of central bank communication affect financial market and real variables. Second, our measure can be used to evaluate the effectiveness of central bank communication including forward guidance. Regardless of whether the BOK intends or not, our indicators based on the minutes help explain the future decision of monetary policy. For example, if we construct a high-frequency MP sentiment indicator before and after the BOK's announcements, it can be used to measure "surprises" around announcements. To measure surprises caused by the Fed's monetary policy announcements, Gertler and Karadi (2015) use the changes in the federal funds futures rate from ten minutes before an announcements to 20 minutes after the announcement. Given that there is no such thing as futures for policy rates in South Korea, a text-based indicator would be a decent alternative. Third, our methodology can be easily applied to construct other indicators to measure macroeconomic uncertainty, public expectation about future monetary policy stance, stock market sentiment, and so on. One can examine how changes in these measures affect asset prices or real variables.

While we well recognize that there are more things to be done in this field, we hope that our study, as a starting point, demonstrates that text mining approach can be a useful addition to the BOK's and researchers' toolbox of analyzing monetary policy and achieving its objectives.

References

Apel, M. and M. B. Grimaldi (2014): "How Informative Are Central Bank Minutes?" *Review of Economics*, 65, 655.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty*," *The Quarterly Journal of Economics*, 131, 1593–1636.

Bennani, H. and M. Neuenkirch (2016): "The (home) bias of European central bankers: new evidence based on speeches," *Applied Economics*, 49, 1114–1131.

Bholat, D., J. Brookes, C. Cai, K. Grundy, and J. Lund (2017): "Sending Firm Messages: Text Mining Letters from PRA Supervisors to Banks and Building Societies They Regulate," *Bank of England Working Paper No. 688.*

Bholat, D., S. Hansen, P. Santos, and C. Schonhardt-Bailey (2015): "Text mining for central banks," *Centre for Central Banking Studies Handbook*, 1–19.

Born, B., M. Ehrmann, and M. Frtzscher (2014): "Central Bank Communication on Financial Stability," *Economic Journal*, 124, 701–734.

Choi, H. and H. Varian (2012): "Predicting the Present with Google Trends," *Special Issue: Selected Papers from the 40th Australian Conference of Economists*, 88, 2–9.

Gentzkow, M., B. T. Kelly, and M. Taddy (2017): "Text As Data," *NBER Working Paper 23276.*

Gertler, M. and P. Karadi (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

Hamilton, W. L., K. Clark, J. Leskovec, and D. Jurafsky (2016): "Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora," *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, 2016, 595–605.

Hansen, S. and M. McMahon (2016): "Shocking language: Understanding the macroeconomic effects of central bank communication," *Journal of International Economics*, 99, S114 – S133, 38th Annual NBER International Seminar on Macroeconomics.

Jegadeesh, N. and D. Wu (2013): "Word power: A new approach for content analysis," *Journal of financial economics.*

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105, 1177–1216.

Kennedy, J. E. (1994): "The information in diffusion indexes for forecasting related economic aggregates," *Economics Letters*, 44, 113–117.

Lee, D.-J., J.-H. Yeon, I.-B. Hwang, and S.-G. Lee (2010): "Kkma: A Tool for Utilizing Sejong Corpus Based on Relational Database," *Journal of KIISE: Computing Practices and Letters*, 16, 1046–1050 [in Korean].

Lee, Y. (2018): "Introduction to eKoNLPy: A Korean NLP Python Library for Economic Analysis," https://github.com/entelecheia/eKoNLPy.

Liu, B. (2009): "Sentiment Analysis and Subjectivity," in *Handbook of Natural Language Processing*, New York, NY, USA: Marcel Dekker, Inc.

Loughran, T. and B. McDonald (2011): "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of Finance*.

Lucca, D. and F. Trebbi (2011): "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements," .

McLaren, N. and R. Shanbhogue (2011): "Using Internet Search Data as Economic Indicators," *Bank of England Quarterly Bulletin*.

Meinusch, A. and P. Tillmann (2017): "Quantitative Easing and Tapering Uncertainty: Evidence from Twitter," *International Journal of Central Banking*.

Nopp, C. and A. Hanbury (2015): "Detecting Risks in the Banking System by Sentimental Analysis," *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*.

Nyman, R., S. Kapadia, D. Tuckett, D. Gregory, P. Ormerod, and R. Smith (2018): "News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment," *Bank of England Staff Working Paper No. 704*.

Picault, M. and T. Renault (2017): "Words are not all created equal: A new measure of ECB communication," *Journal of International Money and Finance*, 79, 136 – 156.

Porter, M. F. (1980): "An algorithm for suffix stripping," *Program*, 14, 130–137.

Pyo, D.-J. and J. Kim (2017): "News Media Sentiment and Asset Prices: Text-mining approach," *KIF Working Paper*.

Shin, M., B. Zhang, M. Zhong, and D. J. Lee (2018): "Measuring international uncertainty: The case of Korea," *Economics Letters*, 162, 22 – 26.

Söderlind, P. and L. Svensson (1997): "New techniques to extract market expectations from financial instruments," *Journal of Monetary Economics*, 40, 383–429.

Stone, P. J., D. C. Dunphy, and M. S. Smith (1966): *The general inquirer: A computer approach to content analysis.*, MIT Press.

Tetlock, P. C. (2007): "Giving content to investor sentiment: The role of media in the stock market," *The Journal of finance*, 62, 1139–1168.

Tobback, E., S. Nardelli, and D. Martens (2016): "Between hawks and doves: measuring central bank communication," 1–41.

Turney, P. D. (2002): "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews," in *Proceedings of the 40th annual meeting on association for computational linguistics*, Association for Computational Linguistics, 417–424.

Warsh, K. (2014): "Transparency and the Bank of England's Monetary Policy Committee," *Review by Kevin Warsh*.

Won, J.-H., W. Son, H. Moon, and H. Lee (2017): "Text mining techniques for classification of economic sentiment," *BOK Quarterly Bulletin*.

Zhao, Z., T. Liu, S. Li, B. Li, and X. Du (2017): "Ngram2vec: Learning Improved Word Representations from Ngram Co-occurrence Statistics," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 244–253.

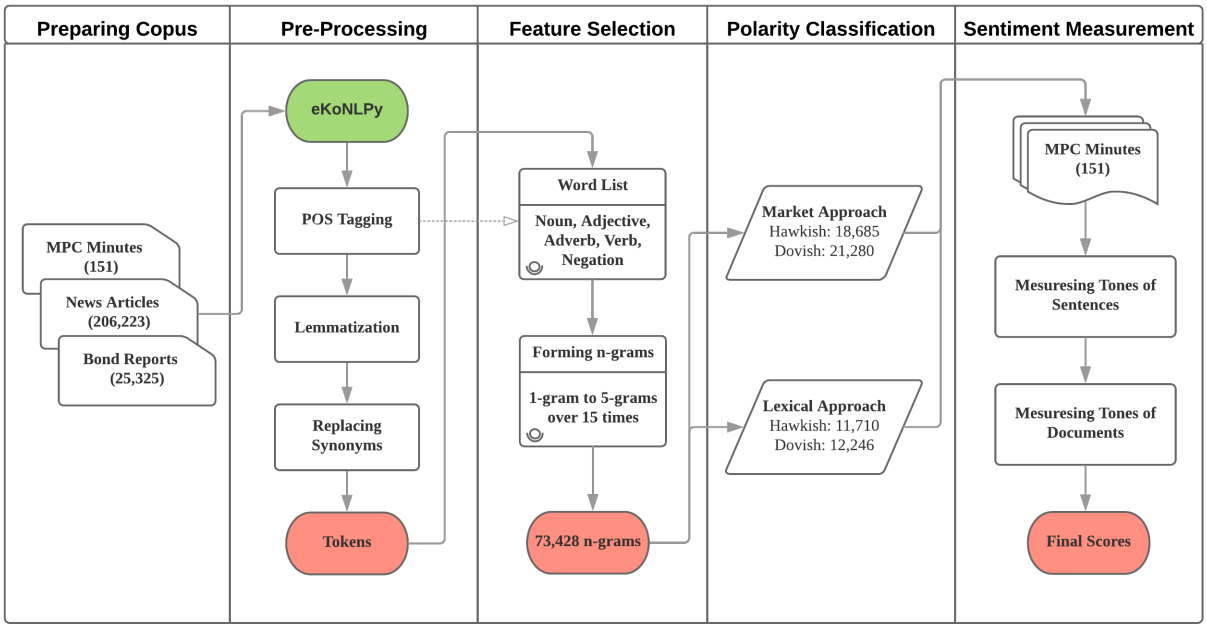
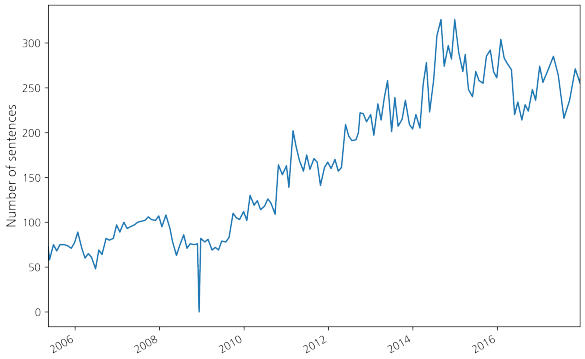
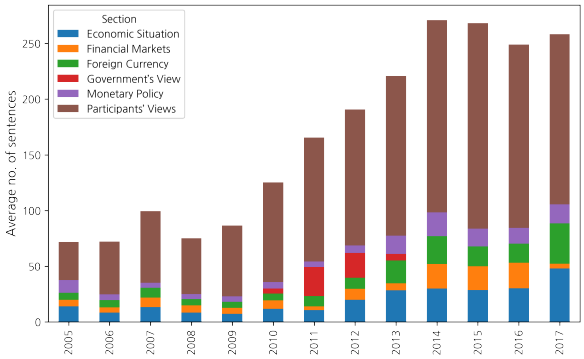


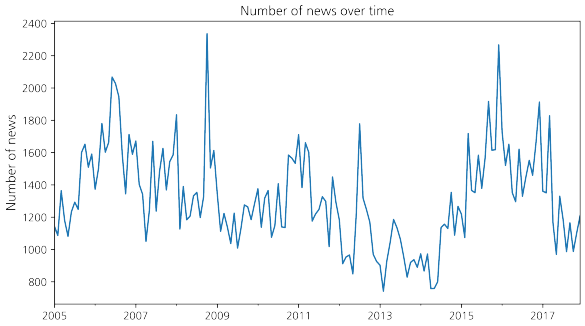
Fig. 1. Procedure of Sentiment Analysis
This figure, along with table 1, summarizes our discussion in section 3.



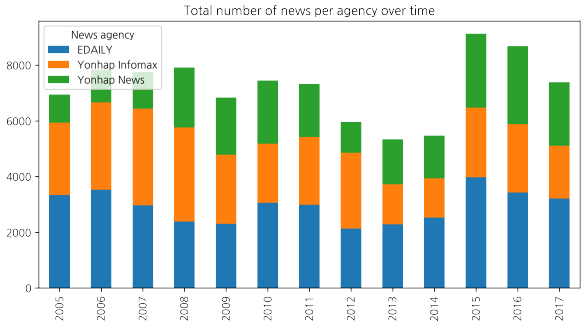
(a) Number of sentences in MPC minutes



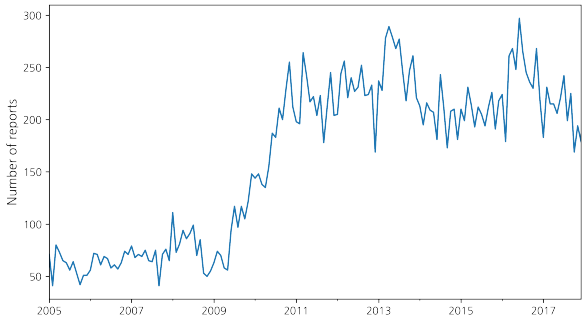
(b) Number of sentences in MPC minutes, by section



(c) Number of news for each month



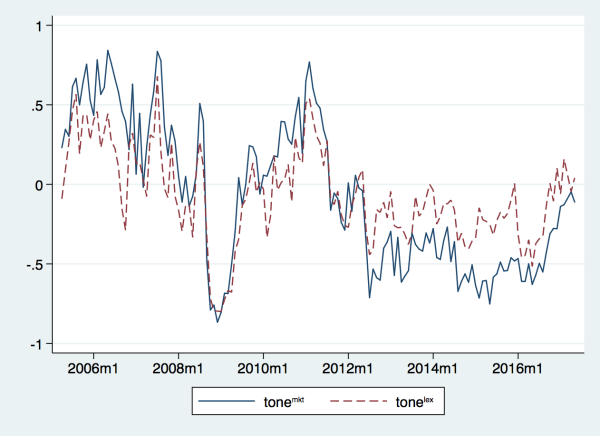
(d) Number of news, by sources



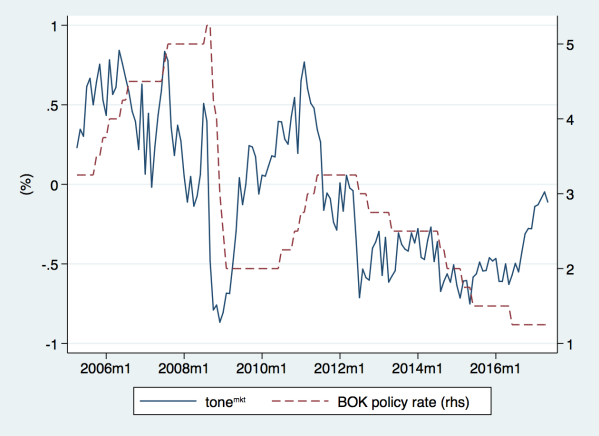
(e) Number of bond analysts' reports

Fig. 2. Text data

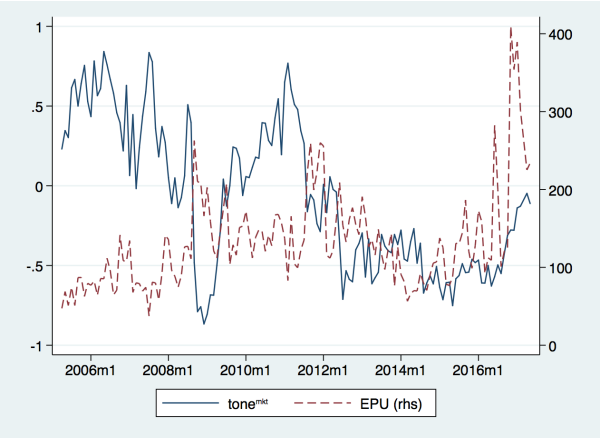
Fig. 3. Topic wordclouds of the corpus



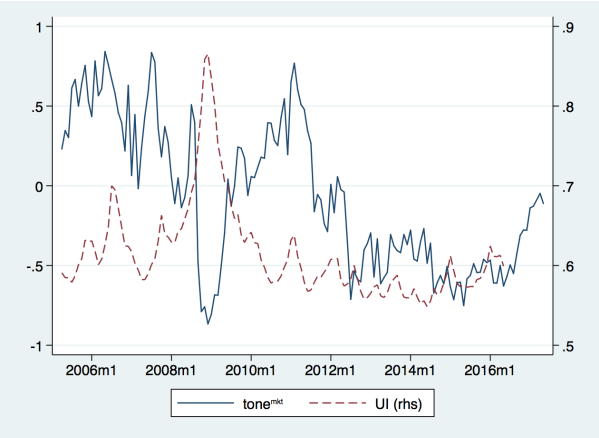
(a) $tone^{mkt}$ and $tone^{lex}$



(b) $tone^{mkt}$ and BOK policy rate

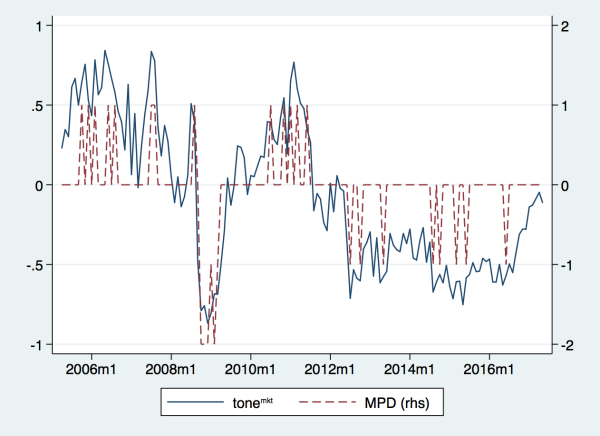


(c) $tone^{mkt}$ and EPU

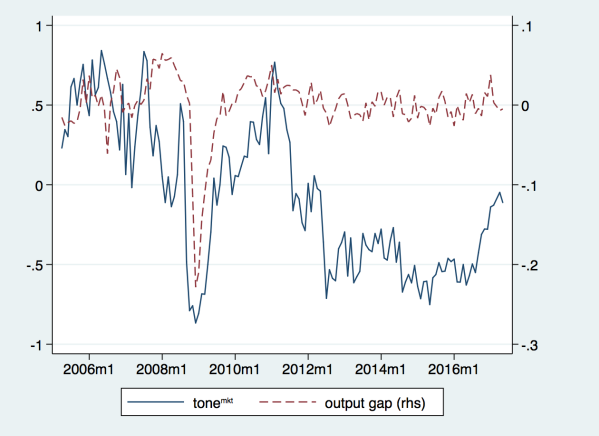


(d) $tone^{mkt}$ and UI

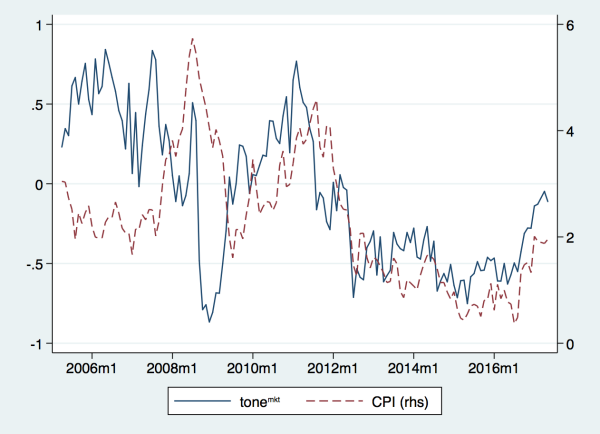
Fig. 4. MP sentiments, BOK policy rate, and other measures of uncertainty
This figure shows the time series of our text-based indicators $tone^{mkt}$ and $tone^{lex}$ with the BOK policy rate, the Korean version of economic policy uncertainty measure by Baker et al. (2016), and the Korean version of macroeconomic uncertainty measure based on Jurado et al. (2015) and Shin et al. (2018).



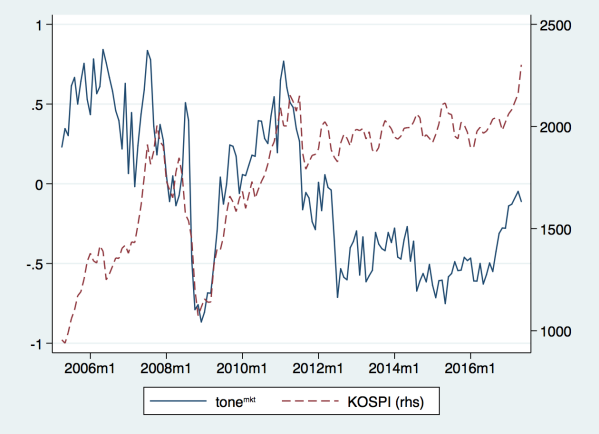
(a) $tone^{mkt}$ and MP decisions



(b) $tone^{mkt}$ and output gap

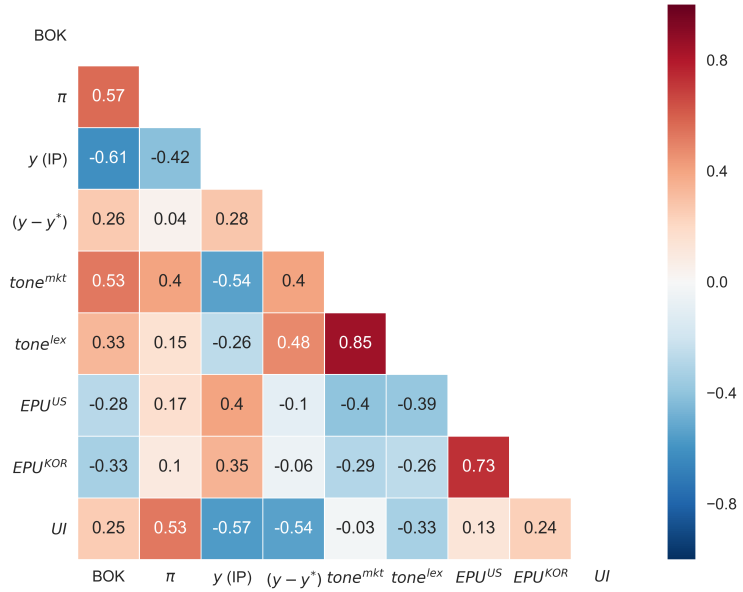


(c) $tone^{mkt}$ and CPI

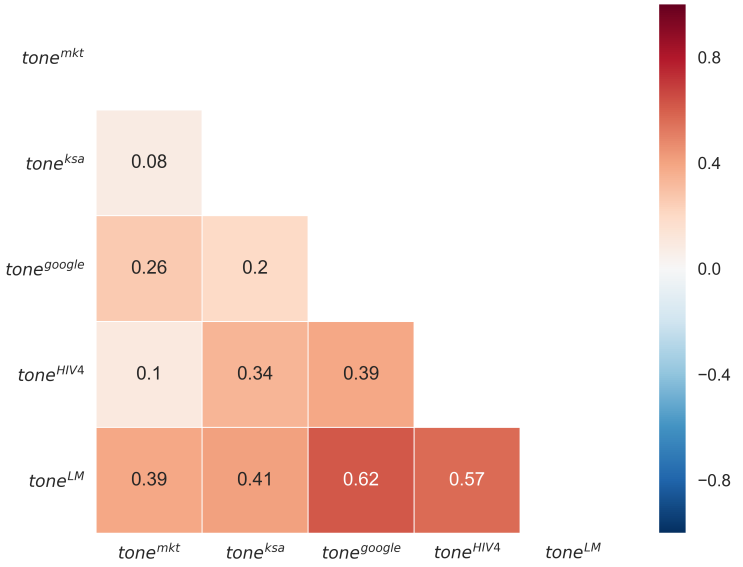


(d) $tone^{mkt}$ and stock market index

Fig. 5. MP sentiment and macroeconomic variables
Panel (a)–(d) show the time series of $tone^{mkt}$, monetary policy decision MPD_t, output gap, CPI, and stock market index (KOSPI).



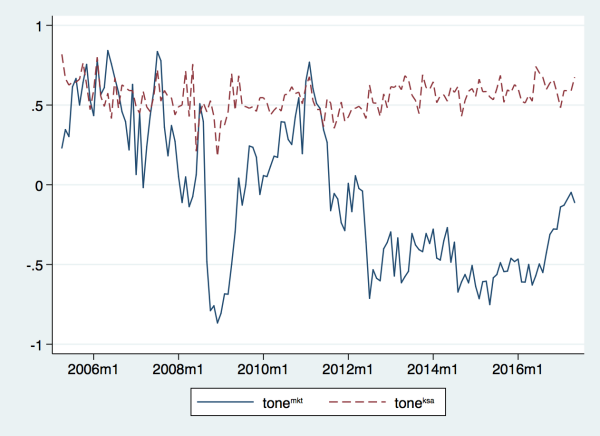
(a) With macroeconomic variables



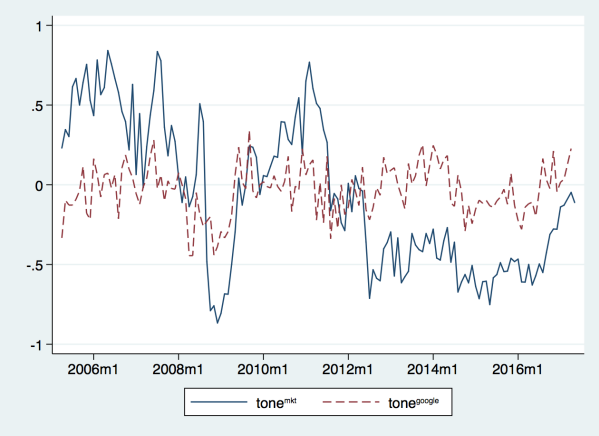
(b) With other text-based indicators

Fig. 6. Correlation coefficients

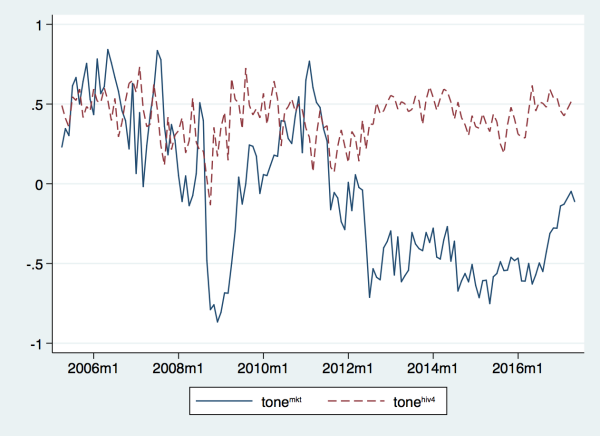
Panel (a) and (b) show the correlation coefficients among key variables. Among the correlation coefficients with $tone_t^{mkt}$, $Corr(tone_t^{mkt}, UI_t)$ $Corr(tone_t^{mkt}, tone_t^{ksa})$, and $Corr(tone_t^{mkt}, tone_t^{HIV4})$ are not statistically significant at 10% level.



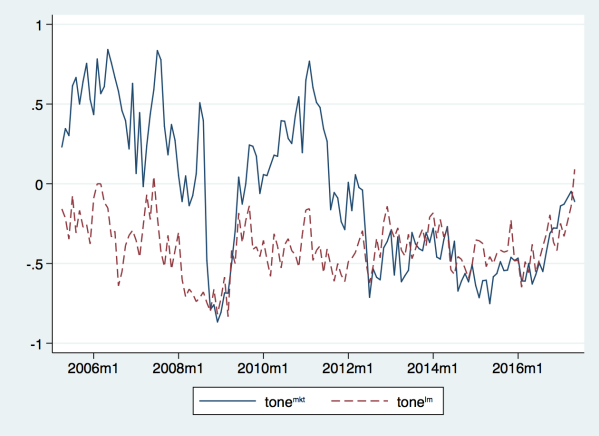
(a) $tone^{mkt}$ and $tone^{ksa}$



(b) $tone^{mkt}$ and $tone^{google}$



(c) $tone^{mkt}$ and $tone^{HIV4}$



(d) $tone^{mkt}$ and $tone^{LM}$

Fig. 7. MP sentiment and other text-based indicators
Panel (a)–(d) show the time-series of text-based indicators. Correlation coefficient of $tone^{mkt}$ with $tone^{ksa}$, $tone^{google}$, $tone^{HIV4}$, and $tone^{LM}$ are 0.08, 0.26, 0.10, and 0.40, respectively.

Table 1: Process of sentiment analysis
This table summarizes what we do in each step of sentimental analysis. See section 3 for more details.

1. Preparing the corpus	• 231,699 documents – 151 MPC minutes – 206,223 news articles related to interest rates – 25,325 bond analyst reports
2. Pre-processing texts	• Tokenization • Normalization – removing stop words – stemming and lemmatization • Morphological analysis of the Korean language → eKoNLPy – spacing – terminology and foreign words
3. Feature selection	• N-grams as a feature – 73,428 n-grams
4. Polarity classification	• Market approach – Naive-Bayes classifier • Lexical approach – ngram2vec – SentProp of Hamilton et al. (2016) • Evaluation – 2,341 manually classified sentences – out-of-sample test
5. Sentiment measurement	• Measuring tones of sentences from the features of n-grams • Measuring tones of documents from tones of sentences

Table 2: Statistics of the corpus

Document type	No. of docs	Average no. of senteneces	Max no. of senteneces
MPC minutes	151	165	326
News articles	206,223	15	340
Bond reports	25,325	49	2,515
Total	231,699	19	2,515

Table 3: Average weights of the topics

No.	Topic name	Total	Minutes	News	Report
1	Foreign Currency	5.24	11.20	5.94	3.75
2	Financial Policy	2.69	2.24	3.15	1.99
3	Bond Issue Market 1	3.35	0.73	1.29	6.79
4	Monetary Policy	3.81	12.56	4.47	2.20
5	Bond Issue Market 2	2.79	1.32	2.67	3.08
6	Financial Crisis	1.79	1.03	2.09	1.36
7	Swap Market	4.30	3.05	4.02	4.82
8	Inflation	3.32	10.56	2.68	3.89
9	Credit Ratings	1.42	0.38	0.93	2.26
10	Real Estate	1.20	0.15	1.71	0.46
11	Global Gvrnmt Bond	2.23	0.40	1.34	3.78
12	Macro Stability	1.26	2.74	1.16	1.32
13	Real Estate Policy	3.01	2.44	3.99	1.46
14	Eurozone	3.94	1.27	3.33	5.07
15	Economic Growth	5.02	13.60	4.58	5.19
16	Money Market	0.88	0.79	1.09	0.55
17	Global Stock Market	2.11	0.37	2.74	1.20
18	Global Monetary Policy	5.52	4.87	4.36	7.40
19	Financial Instruments	2.34	0.22	3.42	0.74
20	Corporate Valuation	1.64	0.27	2.13	0.95
21	Capital Requirements	0.77	0.42	0.48	1.27
22	Gvrnmt Bond Futures	4.69	0.46	4.44	5.34
23	Domestic Stock Market	0.81	0.17	1.21	0.23
24	Soho Money Market	1.34	0.65	2.06	0.22
25	Bond Market Price	1.01	0.33	0.52	1.82
26	Carry Trade	1.41	0.66	1.87	0.71
27	Gvrnmt Regulation	1.43	0.87	1.75	0.95
28	Fund Market	4.53	2.19	5.40	3.30
29	Corporate Restructuring	1.43	1.66	1.59	1.16
30	Comsumption/Income	1.93	5.68	1.72	2.04
31	Liquidity Provision	0.61	0.21	0.46	0.88
32	Politics	2.13	0.65	1.51	3.21
33	Raw Material Price	1.61	0.90	1.54	1.78
34	Bond Investment Strategy	1.80	0.13	1.44	2.46
35	Housing Debt	5.99	8.03	7.81	2.95
36	Corporate Credit Ratings	3.61	0.46	1.86	6.60

Table 4: A sample of polarity lexicon by market approach

Hawkish	Dovish
총액/NNG;대출/NNG;한도/NNG;축소/NNG	금리/NNG;지준율/NNG;인하/NNG
재할인율/NNG;인상/NNG	하락/NNG;거래/NNG;마치/VV;내리/VV
인플레이션/NNG;심리/NNG;확산/NNG	금리/NNG;인하/NNG;가계/NNG;부채/NNG;증가/NNG
자본/NNG;유출입/NNG;변동/NNG;완화/NNG	금리/NNG;인하/NNG;소식/NNG;상승/NNG
자본/NNG;유출입/NNG;규제/NNG;우려/NNG	정책/NNG;공조/NNG;금리/NNG;인하/NNG
물가/NNG;안정/NNG;견조/NNG;성장/NNG	실물/NNG;경기/NNG;침체/NNG;우려/NNG
금리/NNG;인상/NNG;인플레이션/NNG;우려/NNG	유동성/NNG;경색/NNG;해소/NNG
소비자/NNG;물가/NNG;상승률/NNG;금리/NNG;인상/NNG	경제주체/NNG;심리/NNG;부진/NNG
물가/NNG;불안/NNG;확산/NNG	금융위기/NNG;세계/NNG;확산/NNG
잠재/NNG;성장률/NNG;경제/NNG;성장/NNG	구조조정/NNG;자본/NNG;확충/NNG
완만/NNG;속도/NNG;확장/NNG	비둘기/NNG;금리/NNG;인하/NNG
콜/NNG;금리/NNG;인상/NNG;금리/NNG;인상/NNG	위기설/NNG;불안/NNG
총액/NNG;한도/NNG;대출/NNG;금리/NNG;인상/NNG	인하/NNG;파급효과/NNG
금리/NNG;인상/NNG;물가/NNG;상승/NNG;압력/NNG	유로존/NNG;경제/NNG;지표/NNG;부진/NNG
경제/NNG;예상/NNG;회복/NNG	경기/NNG;후퇴/NNG;우려/NNG;완화/NNG
금리/NNG;인상/NNG;콜/NNG;금리/NNG;인상/NNG	신용스프레드/NNG;부담/NNG
발행/NNG;압력/NNG;약화/NNG	국고채/NNG;하락/NNG;금리/NNG;내리/VV
자본/NNG;규제/NNG;우려/NNG	금리/NNG;인하/NNG;지준율/NNG;인하/NNG
금리/NNG;콜/NNG;인상/NNG	금리/NNG;인하/NNG;경계/NNG
금리/NNG;인상/NNG;물가/NNG;상승/NNG	정책/NNG;공조/NNG;차원/NNG;금리/NNG;인하/NNG
자금/NNG;해외/NNG;이탈/NNG	금리/NNG;인하/NNG;효과/NNG;없/VA
경기/NNG;위축/NNG;속도/NNG;둔화/NNG	금리/NNG;인하/NNG;실망/NNG
물가/NNG;상승/NNG;압력/NNG;점차/MAG;크/VV	미국발/NNG;금융/NNG;불안/NNG
자산시장/NNG;불안/NNG	cp/NNG;금리/NNG;급락/NNG
세제개편안/NNG;불확실성/NNG	cd/NNG;금리/NNG;인하/NNG
금리/NNG;인상/NNG;물가/NNG;불안/NNG	대출/NNG;예금/NNG;금리/NNG;인하/NNG
금리/NNG;인상/NNG;소식/NNG;상승/NNG	금융위기/NNG;세계/NNG;경제/NNG;침체/NNG
원화/NNG;절상/NNG;금리/NNG;인상/NNG	경제/NNG;구조적/VAX;취약/NNG
고유가/NNG;높/VA	금리/NNG;예금/NNG;금리/NNG;내리/VV
금리/NNG;인상/NNG;경기/NNG;위축/NNG	경제/NNG;수출/NNG;감소/NNG

Table 5: Seed words for polarity induction

Positive		Negative	
높/VA	팽창/NNG	낮/VA	축소/NNG
인상/NNG	매파/NNG	인하/NNG	비둘기/NNG
성장/NNG	투기/NNG;억제/NNG	둔화/NNG	악화/NNG
상승/NNG	인플레이션/NNG;압력/NNG	하락/NNG	회복/NNG;못하/VX
증가/NNG	위험/NNG;선호/NNG	감소/NNG	위험/NNG;회피/NNG
상회/NNG	물가/NNG;상승/NNG	하회/NNG	물가/NNG;하락/NNG
과열/NNG	금리/NNG;상승/NNG	위축/NNG	금리/NNG;하락/NNG
확장/NNG	상방/NNG;압력/NNG	침체/NNG	하방/NNG;압력/NNG
긴축/NNG	변동성/NNG;감소/NNG	완화/NNG	변동성/NNG;확대/NNG
흑자/NNG	채권/NNG;가격/NNG;하락/NNG	적자/NNG	채권/NNG;가격/NNG;상승/NNG
견조/NNG	요금/NNG;인상/NNG	부진/NNG	요금/NNG;인하/NNG
낙관/NNG	부동산/NNG;가격/NNG;상승/NNG	비관/NNG	부동산/NNG;가격/NNG;하락/NNG
상향/NNG	(Total 25 seeds)	하향/NNG	(Total 25 seeds)

Table 6: A sample of polarity lexicon by lexical approach

Hawkish	Dovish
인상/NNG	인하/NNG
확장/NNG	하향/NNG
상향/NNG	부진/NNG
투기/NNG;억제/NNG	회복/NNG;못하/VX
금리/NNG;상승/NNG	금리/NNG;하락/NNG
상회/NNG	악화/NNG
채권/NNG;가격/NNG;하락/NNG	침체/NNG
인플레이션/NNG;압력/NNG	하락/NNG
과열/NNG	변동성/NNG;확대/NNG
견조/NNG	위축/NNG
팽창/NNG	하회/NNG
물가/NNG;상승/NNG	둔화/NNG
부동산/NNG;가격/NNG;상승/NNG	완화/NNG
성장/NNG	채권/NNG;가격/NNG;상승/NNG
긴축/NNG	물가/NNG;하락/NNG
흑자/NNG	위험/NNG;회피/NNG
요금/NNG;인상/NNG	하방/NNG;압력/NNG
상방/NNG;압력/NNG	부동산/NNG;가격/NNG;하락/NNG
낙관/NNG	비관/NNG
변동성/NNG;감소/NNG	요금/NNG;인하/NNG
위험/NNG;선호/NNG	적자/NNG
매파/NNG	비둘기/NNG
부동산/NNG;과열/NNG;억제/NNG	둔화/NNG;경기/NNG;침체/NNG
부동산/NNG;과열/NNG	경기/NNG;침체/NNG;빠지/VV
과열/NNG;우려/NNG	악화/NNG;경기/NNG;침체/NNG
과열/NNG;억제/NNG	경기/NNG;침체/NNG
과열/NNG;막/VV	침체/NNG;빠지/VV
경기/NNG;과열/NNG	침체/NNG;가능성/NNG;높/VA
부동산/NNG;과열/NNG;우려/NNG	경기/NNG;침체국면/NNG;빠지/VV
경기/NNG;과열/NNG;우려/NNG	침체/NNG;경기/NNG;침체/NNG
가격/NNG;억제/NNG	둔화/NNG;침체/NNG
투자/NNG;과열/NNG	경기/NNG;침체/NNG;빠지/VV;않/VX
부동산/NNG;가격/NNG;억제/NNG	이미/MAG;침체/NNG
경기/NNG;과열/NNG;억제/NNG	길/VA;침체/NNG
과열/NNG;조짐/NNG	침체/NNG;빠지/VV;우려/NNG
인플레이션/NNG;긴축/NNG	침체국면/NNG;빠지/VV
경기/NNG;과열/NNG;막/VV	이미/MAG;경기/NNG;침체/NNG
경제/NNG;과열/NNG	침체/NNG;최악/NNG
긴축/NNG;압력/NNG	경제/NNG;침체/NNG;빠지/VV
과열/NNG;방지/NNG	침체/NNG;늪/NNG

Table 7: Ordered probit, changes in BOK policy rate

This table displays the ordered probit estimation result of equation (5) and (6) when changes in policy rate ΔBOK are used as changes in monetary policy stance ΔMP. *, **, and *** denote p-value < 0.05, p-value < 0.01, and p-value < 0.001, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: ΔBOK_t					
ΔBOK_{t-1}	1.893**	1.790**	-0.209	-0.839	1.611*	1.296
	(0.622)	(0.632)	(0.736)	(0.797)	(0.642)	(0.725)
$\Delta(\pi_t - \pi^*)$	0.142	0.0163	-0.364	-0.490	-0.0690	0.0274
	(0.331)	(0.341)	(0.517)	(0.431)	(0.348)	(0.352)
$\Delta(y_t - y^*)$	7.068	5.614	6.025	8.351	5.696	4.803
	(4.362)	(4.634)	(5.298)	(5.160)	(4.660)	(4.764)
$\Delta\pi_t^e$		1.734	1.553	1.635	1.948*	1.883*
		(0.910)	(1.262)	(1.107)	(0.928)	(0.923)
Δy_t^e		0.322	0.153	0.0661	0.294	0.313
		(0.450)	(0.637)	(0.536)	(0.456)	(0.455)
$tone_t^{mkt}$			5.327***			
			(1.114)			
$tone_t^{lex}$				4.515***		
				(0.797)		
EPU_t (Korea)					-0.00374	
					(0.00191)	
UI_t (Korea)						-2.886
						(2.155)
N	143	143	143	143	143	133
pseudo R^2	0.076	0.095	0.446	0.364	0.116	0.107
	Dependent variable: ΔBOK_{t+2}					
ΔBOK_t	2.406***	2.339***	0.773	0.692	2.239**	2.017**
	(0.671)	(0.678)	(0.777)	(0.797)	(0.688)	(0.741)
$\Delta(\pi_t - \pi^*)$	0.359	0.326	0.290	0.174	0.268	0.302
	(0.338)	(0.343)	(0.373)	(0.367)	(0.350)	(0.350)
$\Delta(y_t - y^*)$	5.521	5.190	6.167	6.659	5.241	4.720
	(4.721)	(4.945)	(5.127)	(5.073)	(4.948)	(5.041)
$\Delta\pi_t^e$		0.629	0.607	0.536	0.718	0.781
		(0.901)	(0.963)	(0.946)	(0.908)	(0.907)
Δy_t^e		0.0952	0.160	0.00654	0.0702	0.0901
		(0.449)	(0.482)	(0.468)	(0.451)	(0.451)
$tone_t^{mkt}$			1.406***			
			(0.383)			
$tone_t^{lex}$				1.970***		
				(0.542)		
EPU_t (Korea)					-0.00179	
					(0.00205)	
UI_t (Korea)						-2.106
						(2.121)
N	142	142	142	142	142	134
pseudo R^2	0.110	0.113	0.203	0.189	0.117	0.118

Table 8: Ordered probit, changes in MP stance

This table displays the ordered probit estimation result of equation (5) and (6) when the variable of monetary policy decision MPD is used as changes in monetary policy stance ΔMP. *, **, and *** denote p-value < 0.05, p-value < 0.01, and p-value < 0.001, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: MPD_t					
MPD_{t-1}	0.759***	0.748***	-0.0467	-0.275	0.714**	0.583*
	(0.215)	(0.216)	(0.267)	(0.292)	(0.219)	(0.239)
$\Delta(\pi_t - \pi^*)$	0.109	0.0166	-0.336	-0.513	-0.0655	0.00875
	(0.331)	(0.340)	(0.512)	(0.433)	(0.346)	(0.352)
$\Delta(y_t - y^*)$	7.627	6.101	8.136	11.55*	6.247	5.913
	(4.634)	(4.917)	(5.603)	(5.705)	(4.961)	(5.210)
$\Delta\pi_t^e$		1.393	0.959	1.220	1.576	1.600
		(0.894)	(1.218)	(1.086)	(0.909)	(0.910)
Δy_t^e		0.355	-0.0585	-0.183	0.307	0.298
		(0.441)	(0.621)	(0.531)	(0.447)	(0.446)
$tone_t^{mkt}$			5.464***			
			(1.122)			
$tone_t^{lex}$				4.900***		
				(0.853)		
EPU_t (Korea)					-0.00364	
					(0.00189)	
UI_t (Korea)						-3.829
						(2.105)
N	143	143	143	143	143	133
pseudo R^2	0.095	0.109	0.461	0.397	0.128	0.130
	Dependent variable: MPD_{t+2}					
MPD_{t-1}	0.800***	0.780***	0.131	0.0958	0.737***	0.629**
	(0.215)	(0.217)	(0.272)	(0.278)	(0.221)	(0.235)
$\Delta(\pi_t - \pi^*)$	0.468	0.444	0.409	0.261	0.374	0.395
	(0.338)	(0.343)	(0.375)	(0.369)	(0.350)	(0.351)
$\Delta(y_t - y^*)$	4.871	4.301	6.145	6.981	4.413	3.922
	(4.746)	(4.982)	(5.243)	(5.220)	(4.987)	(5.113)
$\Delta\pi_t^e$		0.512	0.510	0.413	0.614	0.732
		(0.902)	(0.970)	(0.951)	(0.910)	(0.906)
Δy_t^e		0.169	0.235	0.0406	0.135	0.129
		(0.445)	(0.479)	(0.463)	(0.447)	(0.447)
$tone_t^{mkt}$			1.585***			
			(0.418)			
$tone_t^{lex}$				2.258***		
				(0.581)		
EPU_t (Korea)					-0.00218	
					(0.00205)	
UI_t (Korea)						-3.420
						(2.068)
N	142	142	142	142	142	134
pseudo R^2	0.114	0.116	0.215	0.205	0.122	0.132

Table 9: Comparison of text-based indicators

This table displays the ordered probit estimation result of equation (5) and (6). *, **, and *** denote p-value < 0.05, p-value < 0.01, and p-value < 0.001, respectively.

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: ΔBOK_t				
ΔBOK_{t-1}	-0.209	1.730**	1.440*	1.804**	1.422*
	(0.736)	(0.643)	(0.664)	(0.633)	(0.654)
$\Delta(\pi_t - \pi^*)$	-0.364	0.0230	0.0251	-0.0317	-0.121
	(0.517)	(0.341)	(0.344)	(0.352)	(0.352)
$\Delta(y_t - y^*)$	6.025	5.640	5.378	5.671	6.691
	(5.298)	(4.640)	(4.650)	(4.637)	(4.711)
$\Delta\pi_t^e$	1.553	1.658	1.693	1.867*	1.964*
	(1.262)	(0.923)	(0.918)	(0.940)	(0.925)
Δy_t^e	0.153	0.263	0.0715	0.237	-0.163
	(0.637)	(0.465)	(0.474)	(0.474)	(0.497)
$tone_t^{mkt}$	5.327***				
	(1.114)				
$tone_t^{ksa}$		0.576			
		(1.184)			
$tone_t^{google}$			1.458		
			(0.843)		
$tone_t^{HIV4}$				0.492	
				(0.866)	
$tone_t^{LM}$					1.890*
					(0.783)
pseudo R^2	0.446	0.097	0.111	0.097	0.127
	Dependent variable: ΔBOK_{t+2}				
ΔBOK_t	0.773	2.502***	2.158**	2.102**	1.770*
	(0.777)	(0.700)	(0.686)	(0.705)	(0.709)
$\Delta(\pi_t - \pi^*)$	0.290	0.334	0.325	0.127	0.194
	(0.373)	(0.344)	(0.345)	(0.356)	(0.353)
$\Delta(y_t - y^*)$	6.167	5.419	4.702	5.937	6.766
	(5.127)	(4.981)	(4.964)	(5.048)	(5.064)
$\Delta\pi_t^e$	0.607	0.771	0.634	1.430	1.151
	(0.963)	(0.915)	(0.907)	(0.969)	(0.940)
Δy_t^e	0.160	0.232	-0.106	-0.212	-0.441
	(0.482)	(0.467)	(0.473)	(0.472)	(0.498)
$tone_t^{mkt}$	1.406***				
	(0.383)				
$tone_t^{ksa}$		-1.337			
		(1.197)			
$tone_t^{google}$			1.183		
			(0.815)		
$tone_t^{HIV4}$				2.256*	
				(0.935)	
$tone_t^{LM}$					2.311**
					(0.820)
pseudo R^2	0.203	0.119	0.124	0.144	0.156

Appendices

A. BOK non-standard policy announcements

Date	Announcements
October 2008	currency swap with the Fed loan guarantee
November 2008	expansion of Aggregate Credit Ceiling Loans expansion of the range of BOK's counterparties
December 2008	expansion of the range of BOK's counterparties interest on reserves emergency liquidity assistance currency swaps with China and Japan
February 2009	currency swap extension with the Fed
March 2009	expansion of Aggregate Credit Ceiling Loans
June 2009	currency swap extension with the Fed

B. eKoNLPy Tagset for POS Tagging

Table A1: Tagset used in Mecab tagger of eKoNLPy

Tag	Name	Tag	Name
NNG	General Noun	JKQ	Case Postposition (Quotation)
NNP	Proper Noun	JC	Conjunctive Postposition
NNB	General Dependent Noun	JX	Auxiliary Postposition
NNBC	Unit Word	EP	Prefinal Ending
NR	Number Word	EF	Final Ending
NP	Pronoun	EC	Conjunctive Ending
VV	Verb	ETN	Nominal Ending
VA	Adjetive	ETM	Adnominal Ending
VAX	Derived Adjetive	XPN	Noun Prefix
VX	Auxiliary Predicate	XSN	Noun Suffix
VCP	Positive Copula	XSV	Verbalization Suffix
VCN	Negative Copula	XSA	Adjectivization Suffix
MM	Determiner	XR	Root Word
MAG	Adverb	SF	Sentence Ending Marker
MAJ	Conjunctive Adverb	SE	Ellipsis Symbol
IC	Exclamaation	SSO	Left Quotation Mark
JKS	Case Postposition (Nominative)	SSC	Right Quotation Mark
JKC	Case Postposition (Complementive)	SC	Separator Symbol
JKG	Case Postposition (Determinative)	SY	Symbol
JKO	Case Postposition (Objective)	SH	Chinese Charachter
JKB	Adverbial Postposition	SL	Foreign Word
JKV	Case Postposition (Vocative)	SN	Number