More Than Words:
A Textual Analysis of Monetary Policy Communication

Pongsak Luangaram[1]

Warapong Wongwachara[2]

September 2016

Abstract

This paper employs various tools from computational linguistics to monetary policy statements to obtain some stylized facts and gain exploratory insights into the nature of central bank communication. The sample was taken from a wide array of central banks, covering major central banks and others under the inflation-targeting (IT) regime in both developed and emerging markets. In total, our data comprise 3,011 entries, spanning from 2000 to 2015. Three major aspects of communication were examined in this study, namely (i) *readability* – the ease with which a reader can understand a written text, (ii) *topics* – the key themes that are discussed in the policy statements, and (iii) *tones* – how positive/negative the outlook is in the central bank's language assessment.

We find that understanding central bank language generally required an advanced reader, although the statements varied substantially across banks in terms of both syntax and structure. The use of academic and unfamiliar words was among main contributors to such complexity. We also noticed that readability tended to fall when central banks lowered policy interest rate. By employing hierarchical clustering analysis to construct the structure of central bank rhetoric, monetary policy deliberation was indeed a complicated decision making in which the authorities must strike a balance between a fairly large number of critical issues. Our topic modelling also reveals that IT central banks generally had communicated more about inflation and price level than economic growth. Moreover, global developments seemed to have carried larger weights in policy deliberation since the breakout of the global financial crisis in 2008. On tone analysis, the Federal Open Market Committee's communication tone on the real economy was consistent with their interest rate decisions, and broadly in line with actual GDP growth. The tone of the ECB's Governing Council, on the other hand, was predictive of both growth and inflation outlook.

Keywords: Monetary policy; Communication; Computational linguistics

JEL classification: E52, E58

---

[1] Faculty of Economics, Chulalongkorn University. Corresponding author: pluangaram@gmail.com.
[2] Monetary Policy Group, Bank of Thailand.

Acknowledgements: We are grateful to Michael McMahon, Piti Disyatat, Yuthana Setthapramote and seminar participants at PIER research exchange for fruitful discussions. Special thanks go to research assistance of Parinya Nimsuwan, Aphichat Wisitkitchakarn, Chalisa Kallayanamitra, and Kanokporn Asanittong. The views expressed in this paper do not necessarily reflect the views of the Bank of Thailand, and any errors in this paper are our own.

More Than Words: A Textual Analysis of Monetary Policy Communication

"I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science, whatever the matter may be."

Baron William Thomson Kelvin

From lecture to the Institution of Civil Engineers, London (3 May 1883), 'Electrical Units of Measurement', Popular Lectures and Addresses (1889), Vol. 1, 80-81.

1. Introduction

Effective communication is now widely recognized as key for maintaining central bank's credibility and ensuring a successful conduct of monetary policy (see e.g. Ehrmann and Fratzscher, 2005; Blinder et al., 2008). The issue is particularly pressing for major central banks, including the Federal Reserve (FED), European Central Bank (ECB) and Bank of Japan (BOJ), whose every single word is scrutinized by market participants. Since the Global Financial Crisis (GFC) in 2008, many central banks have exhausted their traditional policy instruments. Consequently, they resorted to unconventional measures, such as large scale asset purchases, and communication, including giving 'forward guidance', as a mean to strengthen the conduct of monetary policy. Furthermore, for inflation targeting (IT) central banks, clear and consistent communication is of utmost importance in maintaining their transparency and accountability.

Examining the language of central banks, however, is not an easy task. As a remark about monetary policy communication by Praet (2014) exemplifies, 'one has to find the right vocabulary to communicate to a wide audience notions and content that are not immediately intuitive but highly relevant for the lives of many people'. The content of central bank messages does not bode well with clarity either. A policy statement generally implies a vast array of information: assessment of the state of the economy (and its position in the business cycles), analysis of monetary conditions, the stance of monetary policy, and possibly the guidance about future policy, to name but a few.

Recently, tools from computational linguistics have gradually gained popularity among economists – for objectivity and quantifiability – in their attempt to understand central bank communication. The Bank of England (2015), for instance, noted in their One Bank Research Agenda the potential use of textual information and text sources that could help improve the understanding of economic and financial systems. Indeed, qualitative data which are non-numeric such as text and video fit in a standard definition of the so-called 'big data' (Bholat, 2014).[3] A survey of central banks by the Irving Fisher Committee (2015) found a strong interest in big data in central banking community, particularly at senior policy level, and their potential use in conducting central bank policies, as well as supporting macroeconomic and financial stability analyses. Hughes and Kesting (2014), by reviewing works on central bank communication published between 2004-2013, also provide support of linguistic analysis in this area.

[3] Bholat et al. (2015) provide a good background reading on how text mining can be useful for central bank research and policy-making, and basic methods of analysing text as data.

Empirical studies of central bank communication, nevertheless, are often restricted to one or two dimensions of communication apiece, employing one or two computational linguistic measures at most per dimension. The picture of central bank communication documented in the literature is hence far from being complete. More importantly, the works are clustered around communication of the major central banks only, largely leaving the others "smaller" banks unexplored.[4] We notice, for instance, that communication of IT central banks, especially in Emerging Market Economies (EMEs) seems to be much more diverse in terms of both content and style than their counterparts in Advanced Economies (AEs). For example, while AE central banks tend to employ a well-structured and rather repetitive "template" when communicating a policy decision, statements from EMEs may change drastically from one policy meeting to the next. This practice certainly has implications on the effectiveness of monetary policy communication, as well as on the overall monetary policy framework.

In this paper, we examined monetary policy communication of 22 central banks during 2000-2015, by applying advanced computational linguistic tools to their post-meeting monetary policy statements. Our panel data consist of a wide array of central banks, covering major central banks and others under IT in both AEs and EMEs. Our aims are to understand their commonalities and differences, and to obtain stylized facts about and exploratory insights into central bank communication internationally. We measured three main aspects of policy communication, namely (i) *readability* – the ease with which a reader can understand a monetary policy statement, (ii) *topics* – the key areas that are discussed in the policy statements such as growth and inflation, and (iii) *tones* – how positive the economic outlook is in the central banks' assessment.

Our results affirmed a complicated nature of central bank communication. Although the policy statements have gradually become more readable in terms of syntactic structure, the increasing use of academic and unfamiliar words has added to the overall level of complexity. By applying Latent Dirichlet Allocation (LDA) and dictionary methods, we extracted the structural content of central banks' policy statements. Our clustering analysis on the structure of central bank rhetoric showed that IT central banks had been communicating more about inflation and price level than economic growth. On tone analysis, we found that central bank communication has broadly been a mirror of economic and inflation developments, and in some cases, it was predictive of those developments too. On a whole, we believe we have pushed forward the frontier of this vibrant strand of literature.

The rest of the paper is organized as follows. Section 2 describes the data to which we applied computational linguistic tools. Section 3 assesses *readability* of these statements. We pay attention to both time-series and cross-sectional variation in this aspect. Sections 4 and 5 consider *topics* and *tones* of central banks communication, respectively. Each section also discusses previous studies which are relevant to our current work. In the last analysis section, we also present case studies for FED, ECB, and the Bank of Thailand in order to discuss in depth their monetary policy communication. Finally, we conclude, make recommendations to improve the effectiveness of communication.

[4] This is not completely surprising though as major central banks have relied more on communication, as an additional policy instrument, as the GFC.

2. Data

Our raw data are policy statements from 22 central banks obtained from their public websites. The sample contains three major central banks (FED, ECB, and BOJ); the rest are inflation targeters (ITs) from both AEs and EMEs. We consider all available statements during 2000-2015, which spans well over pre- and post-GFC periods. For standardization, we looked at the English version of the statements only. In conducting econometric analyses, additional data for the policy interest rates, GDP growth, inflation, and unemployment rates were collected from CEIC. See Appendix for detailed data descriptions.

The statement data form an unbalanced panel since not all cross-sectional entries (central banks' statements) are observed at the same points in time. This is because meeting dates vary across banks, so do the total numbers of meetings in a year. In total, our data comprise 3,011 observations with a reasonable size of cross-sections (N) and a relatively long time dimension (T), grossing over more than 1.2 million words.

We pre-processed the central bank statements by excluding their pre-ambles, as well as other administrative details that appear on the pages, and then focused only on the corpus of the text. The work files are saved in *.txt* format for carrying out textual analysis. To facilitate econometric analyses, we converted raw data in their original frequencies into a common quarterly basis, by taking average values. For example, ECB's quarterly data are averages of two meetings, while Hungarian National Bank's are averages of three. BOJ's are averages of two-to-four meetings, on the contrary. The transformation compresses the number of observations by about two thirds, but it allows the data to be handled in a standard manner.

3. Readability

We begin our analysis by examining readability – which is generally accepted as a desirable feature of written communication. For one thing, clarity seems to play a key role in enhancing the effectiveness of monetary policy (Levin, 2014). Firstly, we provide a brief literature review on readability in general before narrowing down to previous research on central bank communication. Then we discuss the main tools used in our studies, and present the empirical results accordingly.

3.1 Literature review: Classics and application in central banking

Sherman (1893) is arguably the inaugural work on readability which launched the trend of computational linguistics. He proposed the application of statistical analysis in the study of English literature. His key findings include: Shorter sentences and concrete terms increase readability, whereas spoken language is more efficient than written language. Sherman also made a prediction that written language would become more like spoken language over time.

Subsequent studies started to come up with "formulae" for measuring the degree of readability. Vogel and Washburne (1928) developed an equation that predicts a reading grade of the Stanford Achievement Test:

$$X1 = .085X2 + .101X3 + .604X4 - .411X5 + 17.43,$$

where X1 is the reading score, X2 the number of different words in 1,000, X3 the number of prepositions in 1,000 words, X4 the number of uncommon words in 1,000, and X5 the number of simple sentences in 75.

Lorge's Index (1944, 1948) had fewer elements in his formula: Average sentence length (in words); number of prepositional phrases per 100 words; and number of words not on the Dale list of 769 easy words. As DuBay (2006) argued, Lorge's work established the principles for the readability research that would follow and set the stage for the Flesch Reading Ease formulae (Flesch, 1948).

The formula for the updated Flesch Reading Ease score, one of the mostly cited formula for readability, is as follows.

$$\text{Score} = 206.835 - (1.015 \text{ x ASL}) - (84.6 \text{ x ASW})$$

where: Score = position on a scale of 0 (difficult) to 100 (easy), with 30 = very difficult and 70 = suitable for adult audiences. ASL = average sentence length (the number of words divided by the number of sentences). ASW = average number of syllables per word (the number of syllables divided by the number of words).

Kincaid et al. (1975), based on the works of Flesch's, proposed a formula that calculates a grade level which the passage being evaluated is suitable for.

$$\text{Grade Level} = .39 \text{ (words/sentence)} + 11.8 \text{ (syllables/word)} - 15.59.$$

This is known as the Flesch-Kincaid grade (FK). A sixth-grade level, for instance, is suitable for a person under formal education who is 11-12 years of age. FK is probably the most widely cited measure of readability in the literature on central bank communication.

All of the measures cited above attribute the degree of readability to (i) syntactic structure, and (ii) vocabulary. This is also common among previous works on linguistic analysis of central bank communication. Jansen (2011) applied the Flesch score and the FK grade to the testimony by the FED Chairman at Congressional Monetary Policy Oversight hearings in order to determine the effects of clarity on financial markets. He found that clarity could potentially reduce volatilities of financial variables, especially those of medium-term interests rates, though the impact might vary over time. A recent study by Bulíř et al. (2013) examined time variation in readability of seven central banks.[5] Having applied FK to their inflation reports and press statements, the authors found significant idiosyncratic trends across banks, and concluded that a model for clarity of communication should take into account country- and institution-specific factors too.

The above findings on diversity of communication across central banks are broadly consistent with those of Fracasso et al. (2003) who, based on assessment of independent evaluators, found persistent differences in quality of central bank writing. Another survey of communication practices of 10 central banks in AEs by Kedan and Stuart (2014) also found great breadth and diversity in their communication frameworks.[6]

3.2 ETS TextEvaluator

In addition to FK, we made use of TextEvaluator[7], provided by Educational Testing Service in the United States (thereafter 'ETS index'), in examining clarity of central bank communication. The ETS index is fully-automated web tool that determines (i) the overall level of complexity of a reading passage, henceforth 'ETS Complexity Score' and the appropriate grade-level classification; and (ii) eight possible sources of comprehension

[5] They are Banco Central de Chile, Czech National Bank, ECB, National Bank of Poland, Sveriges Riksbank, Bank of Thailand, and Bank of England.

[6] They are FED, ECB, BOJ, Bank of England, Swiss National Bank, Bank of Canada, Sveriges Riksbank, Norges Bank, Reserve Bank of Australia, and Reserve Bank of New Zealand.

[7] https://texteval-pilot.ets.org/TextEvaluator/

difficulty. The technology is based on research work of Sheehan et al. (2010, 2014) and also further applied and validated in Sheehan (2015, 2016).

ETS Complexity scores are reported on a scale of 100 to 2000; larger value indicates greater complexity, and hence less readability. The eight component scores are reported on a scale that ranges from 1 to 100. The analytical framework is shown in Figure Y. Further information about TextEvaluator can be found on the ETS website.

The first component, *Syntactic Complexity (SYNT)*, assesses complexity of via various features, such as average sentence length and average number of modifiers per noun phrase. The next three components measure vocabulary difficulty, namely *Academic Vocabulary (ACAD)*, *Word Unfamiliarity (UNFA)*, and *Concreteness (CONC)* – words that are more concrete are more likely to facilitate comprehension (less difficult). These four component scores resemble the earlier works, such as the Lorge's and the Flesch-Kincaid indices, in that complexity at linguistic units are deemed to be primary determinants of readability.

Figure 3.1: ETS TextEvaluator Framework – Various Dimensions of Complexity

1. Syntactic Complexity (+)	***Sentence structure***	***ETS Complexity Score***
2. Academic Vocabulary (+) 3. Word Unfamiliarity (+) 4. Concreteness (-)	***Vocabulary difficulty***	
5. Lexical Cohesion (-) 6. Interactive/Conversational Style (-) 7. Level of Argumentation (+)	***Connections across ideas***	
8. Degree of Narrativity (-)	***Organisation***	

This figure illustrates the conceptual framework of ETS TextEvaluator. The overall ETS Complexity Score is based on eight Component Scores which measure different aspects of readability. The +/- signs indicate the contribution of each Component Score to the overall score. For instance, higher Syntactic Complexity contributes positively to the overall score.

The next group of scores is associated with connections across ideas that allow the reader to follow the presented concepts throughout the text. These are *Lexical Cohesion (COHE)*, *Interactive / Conversional Style (CONV)*, and *Level of Argumentation (ARGU)*. As Sheehan et al. (2000, 2014) explain, a coherent message usually contains certain observable features such as repeated instances of the same word stem, and explicit connectives. While COHE and CONV aid comprehension, more formal and argumentative discourse which contain, for instance, subordinating concessive phrases ("although", "however", "on the contrary") and synthetic negations ("nor", "neither") requires greater efforts in following the author's line of reasoning.

The final component score is *Degree of Narrativity (NARR)*. It makes use of the number of words found within quotation marks, referential pronouns, and past-tense verbs, all of which are primary features of any written narrative ("story-like").

Alternative resources

There are other readability tools available on-line. Here, we give a brief review of them.

Coh-Metrix[8] is a system for calculating computational and coherence metrics for written and spoken texts. This tool measures several 'easability' dimensions of the text – all of which are similar to ETS TextEvaluator assessment – namely, *Syntactic Simplicity*, *Deep & Referential Cohesion* (the overlap in content words between local sentences or co-references within the text), *Word Concreteness*, and *Narrativity*. This tool, however, does not allow input text of more than 1,000 words.

Reading Maturity Metric (RMM)[9] by Pearson's Knowledge Technologies Group is a measure of the reading difficulty of texts. It is claimed to be able to predict the grade level of a text, on average, about 30% more accurately than traditional readability formulae. This tool utilizes a range of measures, including text structure, syntax, and vocabulary usage which also covers Word Maturity – an application of Latent Semantic Analysis to discern how each of thousands of individual words gradually develop their unique meanings by being encountered in paragraphs entwined with already known words that implicitly define the new words.[10] At the time of writing, unfortunately, the product was unavailable.

ATOS analyser[11] calculates the readability level for text passages. The results include the overall ATOS level, word count, average word length, and average sentence length. No assessment on other sources of complexity is provided though.

Nelson et al. (2011) assessed the capabilities of six test difficulty metrics, including those outlined above. The authors found that generally the results obtained by these computer-based measures showed strong correlation with reference measures (human graders), especially for informational texts. Nevertheless, the authors also noted that the metrics that included the broader range of linguistic and text measures tended to perform better than those that used only word difficulty and sentence length. This provides further grounds for our use of ETS measures.

In measuring readability, we adhered to TextEvaluator Formatting Guidelines of ETS, including manual removal of non-text elements and inclusion of at least one hard return between each paragraph. The free online version limits the length of ratable passages to 1,600 words. Such limitation created 104 missing observations, or 3.5% of the sample. These are mostly statements of the National Bank of Poland's during 2001-2004, and the ECB's during 2007-2001.

3.3 Empirical analysis and results

This section presents key findings of our analysis on readability. We will proceed by proposing a number of stylized facts obtained by the tools described in the previous section. Both cross-sectional differences and time variation will be considered. In the process, we also gathered syntactic information of central bank statements, i.e. statistics on linguistic units such as words and sentences. These data should give us further insights about the amount of information disclosed by policy makers.

[8] *http://csal.gsu.edu/content/coh-metrix-basic-overview*.

[9] *http://readingmaturity.com*.

[10] For details about the scoring technology, visit http://www.aclweb.org/anthology/P/P11/P11-1031.pdf.

[11] http://www.renaissance.com/products/accelerated-reader/atos/atos-analyzer-for-text .

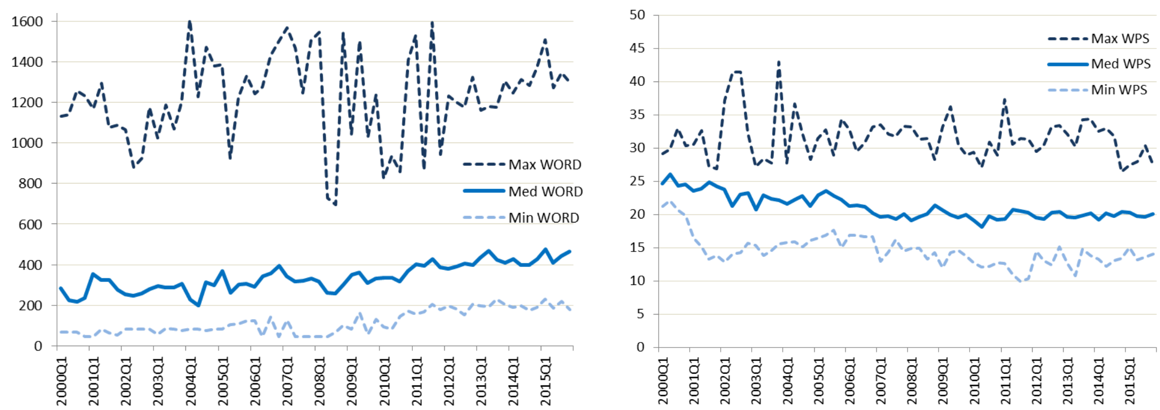
Stylized fact 1: Central banks do talk more over time

Since 2000, the median of total word numbers (WORD) in the statements has been on a rise, from just above 200 words in a statement (half a page) to almost 500 words per statement (a full single page). See Figure 3.2. This finding is consistent with the observation that central banking has subscribed to more transparency and open communication (Blinder et al., 2008). The momentum seems to have accelerated in the aftermath of the financial crisis. We also note positive skewness of the total word number distribution, where central banks who tend to say more are likely to say much more than average.

Despite the increase in WORD, the average number of words per sentence (WPS) has been on a decline. Central banks now use around 5 fewer words in a sentence compared to the past. This is a positive development for readability as it simplifies the sentence structure. We note in passing that the WPS distribution is also less skewed than that of WORD.

Figure 3.2: Basic Statistics of Press Statements

a) Total Words *b) Words per Sentence*



Panel a) shows the median, along with the minimum and maximum, number of total words in a policy statement across 22 central banks over 2000 to 2015. Panel b) shows similar figures for average numbers of words per sentence in a policy statement.

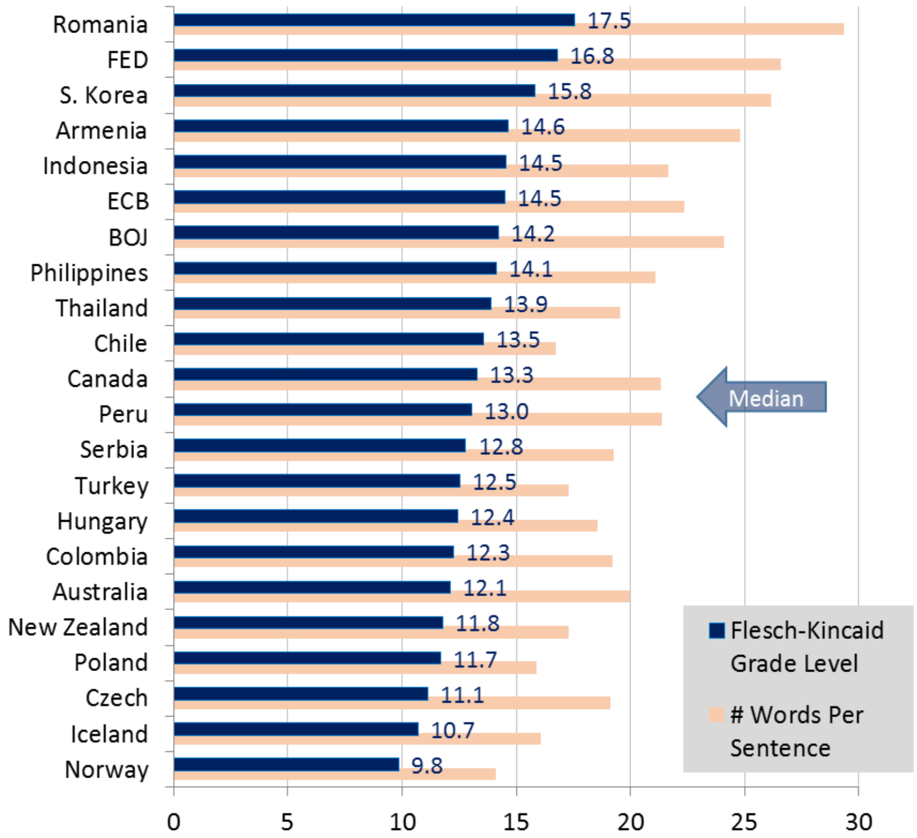
An econometric analysis on our panel data shows that WPS is a very good indicator of readability – lower WPS tends to increase readability. As shown in Table 3.1, WPS has positive and significant sign in both regressions of FK and ETS. WPS alone explains, partly by construction, more than 90% of variation in FK across central banks and time. It also explains about 65% of variation in ETS via contribution to Syntactic complexity.

On an intertemporal basis, both WORD and WPS exhibit large persistence – suggesting that the syntactic complexity of this meeting is determined by that of the previous meeting. This finding might well reflect gradualism in central banking practice where rapid changes are uncommon. In policy communication, this might mean that central banks use / stick to a “template” in statement preparation.

Stylized fact 2: Central bank language requires an advanced reader

Although WPS seems to be lower now than a decade ago, monetary policy statements are far from being light readings. We found that the median FK stayed at around 13 years, implying that central bank statements are suitable for first year undergraduate students or above. See Figure 3.3. All major central banks show above-average FK grades; for example, that of the FED stayed at 16.8 years.

Figure 3.3: Flesch-Kincaid Grade Level Prediction (Avg. 2008-2015)



This figure shows a substantial degree of variation in the Flesch-Kincaid Grade Level Prediction – years of education required in order for a reader to understand the assigned passage – across 22 central banks in our sample. The numbers were 2008-2015 averages as to avoid the supposed structural break during the financial crisis.

In this aspect, FK also shows strong rank-correlation with the ETS measure with Spearman's coefficient of around 85%. The contemporaneous correlation between FK and ETS averages, on the other hand, is around 60%, lower but still considered strong. On a whole, our measures of readability seem to show a satisfactory degree of robustness.

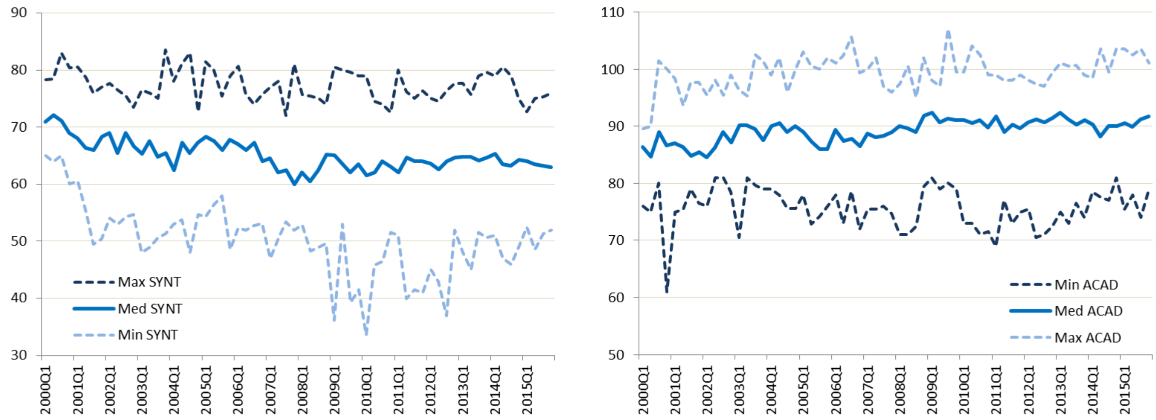
We notice that a couple of central banks which showed low FK were among the top performers on the Dincer-Eichengreen (DE) 'Transparency and Independence' (2014) list, including the Reserve Bank of New Zealand, the Central Bank of Hungary, and the Czech National Bank. Indeed, the average DE score of central banks with below-average FK is larger than the average of those with above-average FK. However, the correlation is not strong. FK does not appear to have any clear distributional pattern across geographical locations or monetary policy regimes either.

Stylized fact 3: Big words contributed a lot to complexity

We looked at the ETS Component scores to gain more insights into different contributors to the overall complexity. In line with the fall in WPS, the median Syntactic complexity score has been on a decline as well. However, this was "compensated" by a rise in Academic words score, resulting in a stable overall ETS Complexity score over our sample.

Figure 3.4: Contributors to ETS Complexity Scores

a) Syntactic complexity score *b) Academic words score*



Panel a) shows the median, along with the minimum and maximum, of ETS Syntactic complexity scores across 22 central banks over 2000 to 2015. Panel b) shows similar figures for Academic words scores.

An attribution analysis, shown in Table 3.1, shows that Syntactic complexity carries the largest weight on the overall complexity score, followed by components that represent vocabulary difficulty, namely Academic vocabulary, Word unfamiliarity, and Concreteness. These weights were not subject to a structural break during the Global Financial Crisis. See Table 3.2. As such, although syntactic features – as similarly measured by FK – that involve understanding sentence remain the main determinants of readability, understanding word seems no less, if not more, important. These are probably the main areas that policy makers should address in order to improve readability of their policy statements.

Our findings probably put into question the current of central bankers, as market participants in general, to use big words and jargons as they raise complexity levels. Economists may consider 'tightening of monetary conditions' a proper description of financial markets in a monetary policy report, while a similar experience is known to traders as 'risk off'. Neither of these makes much sense to a layman though. As Holmes (2014) posits, central bankers have been engaging in communicative experiments that do not merely describe the economy, but actually create its distinctive features.

Table 3.1: Determinants of FK and ETS Complexity Scores

Dep.	FK			ETS				
Eq.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C	6.07	6.04	6.29	1051.07	1018.51	1066.77	38.04	41.42
WPS	0.35	0.35	0.35	11.98	11.37	12.08		
WORD		0.00*			0.10			
SYNT							7.74	7.74
ACAD							5.21	5.20
UNFA							5.63	5.62
CONC							-5.86	-5.87
COHE							-1.31	-1.31
CONV							-1.81	-1.82
ARGU							1.74	1.73
NARR							-0.13	-0.14
RATE			-0.08			-4.65		-0.14
Adj.R-sq	0.91	0.91	0.91	0.65	0.68	0.66	1.00	1.00
# obs.	1,111	1,111	1,055	1,111	1,111	1,055	1,111	1,055

This table shows estimated fixed-effects regressions of readability measures (Flesch-Kincaid and ETS Complexity Scores) on linguistic units (words and words per sentence), complexity characteristics (ETS Component Scores, namely Syntactic, Academic vocabulary, Unfamiliar words, Concreteness, Cohesion, Conversational style, Argumentation, and Narrativity) and the policy interest rate. This set of (unbalanced) panel data include 22 central banks, and span over 64 quarters from Q1 2000 to Q4 2015. Turkey was dropped from (3), (6), and (8) due to unavailability of quarter interest rate data. Estimates of unobserved heterogeneity are omitted. All coefficients are significant 1% or lower level of significance, except those indicated by * which are not significant at 10% level. All specifications reject the null hypothesis of redundant fixed-effects. Adjusted-R sq. of ETS regressions on Component Scores is equal to 1.00 by construction.

Table 3.2: Determinants of FK and ETS Complexity Scores *[Sub-sample Results]*

Dep.	FK						ETS									
Eq.	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Sample	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
C	5.61	5.37	5.59	5.29	5.84	5.45	1050.81	1006.14	1035.21	980.03	1069.99	1027.15	31.91	34.50	31.43	37.00
WPS	0.36	0.39	0.36	0.38	0.37	0.39	11.53	14.48	11.22	12.91	11.81	14.51				
WORD			0.00*	0.00					0.05	0.12						
SYNT													7.84	7.69	7.85	7.69
ACAD													5.20	5.24	5.19	5.23
UNFA													5.63	5.65	5.63	5.65
CONC													-5.68	-5.87	-5.67	-5.87
COHE													-1.41	-1.34	-1.41	-1.35
CONV													-1.80	-1.81	-1.80	-1.81
ARGU													1.72	1.72	1.72	1.71
NARR													-0.11	-0.06	-0.11	-0.06*
RATE					-0.07	-0.03					-5.33	-6.42			0.02*	-0.20
Adj.R-sq	0.93	0.93	0.93	0.93	0.92	0.93	0.74	0.72	0.74	0.75	0.74	0.73	1.00	1.00	1.00	1.00
# obs.	427	684	427	684	404	651	427	684	427	684	404	651	427	684	404	651

This table shows estimated fixed-effects regressions of readability measures (Flesch-Kincaid and ETS Complexity Scores) on linguistic units (words and words per sentence), complexity characteristics (ETS Component Scores, namely Syntactic, Academic vocabulary, Unfamiliar words, Concreteness, Cohesion, Conversational style, Argumentation, and Narrativity) and the policy interest rate. This set of (unbalanced) panel data include 22 central banks, and span over 64 quarters from Q1 2000 to Q4 2015. There are two sub-samples: (a) "Pre-GFC" Q1 2000 – Q4 2007; and (b) "Post-GFC" Q1 2008 – Q4 2015. Turkey was dropped from (3), (6), and (8) due to unavailability of quarter interest rate data. Estimates of unobserved heterogeneity are omitted. All coefficients are significant 1% or lower level of significance, except those indicated by * which are not significant at 10% level. All specifications reject the null hypothesis of redundant fixed-effects. Adjusted-R sq. of ETS regressions on Component Scores is equal to 1.00 by constructi

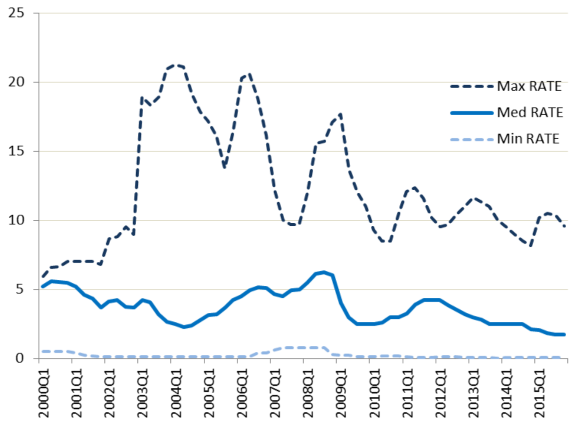
Stylized fact 4: Readability falls during rate cuts

We examined whether interest rate decisions affected readability or not. In order to do so, we first converted individual banks' meeting frequencies into quarterly data such that they can be compared on a standardized basis.

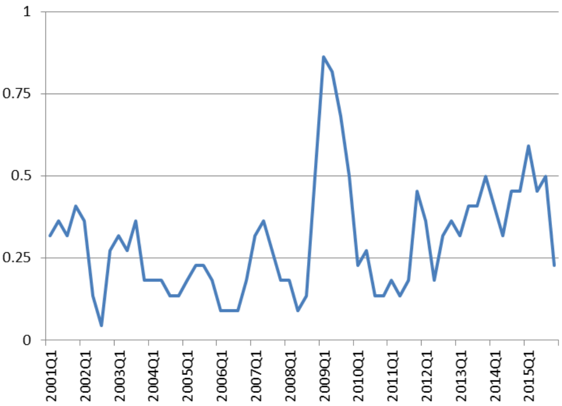
The movements of policy interest rates are shown in Figure 3.5. We saw synchronized movements of central bank decisions during the GFC where most banks in our sample cut their policy interest rate, resulting in the declining median rate. The easing stance continued in the post-GFC period, but this trend went on a halt in 2015 as fewer banks lowered their key rates. This observation well reflects monetary policy divergence that became more prominent in that year.

Figure 3.5: Movements of Policy Interest Rates

a) Rate distribution



b) Proportion of easing banks



Panel a) shows the median, along with the minimum and maximum, of quarterly-average policy interest rates across 21 central banks over 2000 to 2015. Turkey was dropped due to data limitation. Panel b) shows the associated proportion of central banks which lowered policy rates in that quarter.

We found that readability tends to fall when central banks lower their policy interest rates. In Table 3.1, the coefficient of policy interest rate is negative and highly significant in all specifications. Table 3.2 presents a robustness check where we divided our sample into 'Pre-GFC' and 'Post-GFC'. The pervious conclusion still holds more or less.

It is well documented that monetary policy is less effective in shoring up the economy during the downturn, than in slowing it during the economic upturn (see e.g. Mumtaz and Surico, 2015). The above is probably more bad news for central banks when they want to increase the effectiveness of a rate cut – by communicating more clearly and enhancing the expectation channel.

While our statistical analysis does not offer much insight as to why readability tends to fall during easing periods, it could be a result of MPC's attempt to justify a rate cut – which is generally negative for financial stability, though stimulative to growth. For instance, monetary policy easing could potentially hurt commercial banks' profitability in a bad economy where credit risks rise. MPC might, as a result, feel obliged than to communicate more to market participants so that they understand the decisions better. As Holmes (2014) argued, the modern conduct of macroeconomic policies hinge on persuasive communicates aimed at the general public, as well as various groups of stakeholders. Statistically speaking, for every 25bps cut, a statement will contain around 5 more words per statement on average. However, as conventional wisdom goes and also put in the context of monetary policy communication by Dennis and Williams (2007), our results seem to suggest that less is more.

4. Topic and tone analysis of the policy statements

4.1 Text mining methodology

4.1.1 Pre-processing documents

The first pre-processing requirement of turning text into data is to discard sentence structure or order of words in a document. At first sight, one might ask if a sentence structure should be the critical part for understanding the meaning in any textual document. For the purpose of topic discovery, however, we will illustrate later that identifying topics in the monetary policy statements could be made invariant to word order. For example, as shown in the word clouds below: topic about inflation should result in specific terms like consumer, price, core, energy; the topic about economic growth should result in terms like growth, economy, demand, recovery, exports, investment etc. The final output of pre-processing documents involves constructing a 'term-document matrix' which basically containing word frequency that observes in a collection of document.

Since the text corpus is generally very unstructured (e.g. containing number, capital letters, symbols, etc.), the term-document matrix needs to be in a clean matrix form. In addition, we also need to reduce large dimensionality of the matrix. Therefore, the standard procedure usually involves two steps (i.e. cleaning and stemming). First, taking out common words (or stop words) as well as punctuation and capitalization, which appear very often in the text such as 'a', 'an', 'the', 'of' etc. Second, to reduce the total number of unique words in the data set, the second step involves stemming by removing the end of words and counting only stems, for example, the term like inflation becomes 'inflat'. So this stemming process attempts to group words that are grammatically different but thematically identical. The popular stemmer is the Porter algorithm. For more details and illustrated examples of this pre-processing textual document, we refer to Bholat et al. (2015) that gives excellent overview of text mining for economists.

4.1.2 Topic modeling: Latent Dirichlet Allocation (LDA)

Once the document collection has been pre-processed, a statistical model can be implemented to analyze patterns in the data. Specifically, this paper uses topic modeling to simultaneously discover the topics and estimate its proportions in the press statement. Generally, the topic modeling approach assumes that words are an important element for revealing a topical content covered in the text.

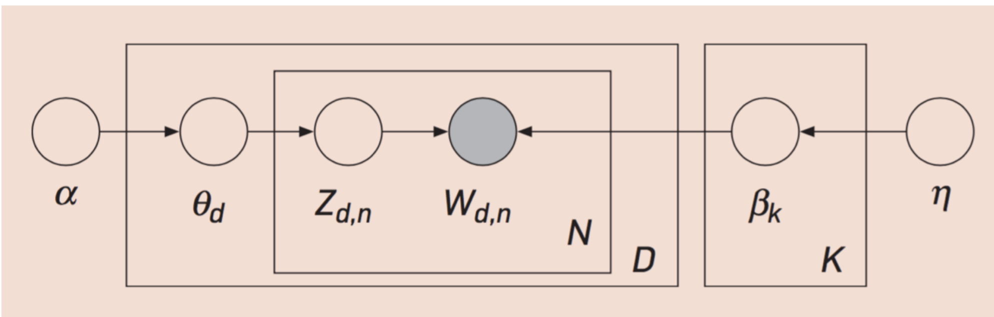
The most popular topic classification model was called Latent Dirichlet Allocation (LDA), firstly introduced by Blei, Ng, and Jordan (2003). The underlying concept is that each document in a corpus is a mixture of an underlying set of topics, i.e. the document is simply a combination of various topics in different proportion. And each topic is a mixture of words with different probabilities. It should be stressed that the LDA algorithm does not need *prior* topic labeling because researcher should be able to interpret by examining collection of keywords in each topic.

As illustrated by Blei (2012), the LDA algorithm involves two steps. First, for each document, the process starts by randomly choosing a distribution over topics. In the second step, each word in each document is drawn from one of the topics where the selected topic is chosen from the per-document distribution over topics. So all the documents will always share the same number of topics with different proportion. While we can observe/read the documents, Blei (2012) note that the goal of topic modeling is to automatically discover hidden topics structure including per-document topic distributions and the per-document per-word topic assignments.

A formal graphical representation can be shown in Figure 4.1, each document d has a multinomial distribution θ_d of topics, and each topic z has a multinomial distribution ϕ_z of words. A document's topic distribution is randomly sampled from a Dirichlet distribution with alpha hyperparameter (α), and each topic's word distribution is randomly sampled from a Dirichlet distribution with beta hyperparameter (β). The alpha represents document-topic density while the beta represents topic-word density. As noted in Hansen, McMahon and Prat (2014), it is recommended to take Griffiths and Steyvers (2004) suggestion by setting the hyperparameter of the Dirichlet prior on topics to $200/V$, where V is the number of unique vocabulary elements, and the hyperparameter of the Dirichlet prior on document-topic distributions to $50/K$, where K is the number of topics.

Thus, topic assignment in LDA is modeled probabilistically and based solely on the document's word token content. In other words, it assumes that meanings associated with a topic can be understood as a set of word clusters, so-called 'bag of words', which captures the probability of word co-occurrence regardless of language syntax or location within a document. A topic can be thought of word clusters that tend to come up in a discussion and therefore co-occur more frequently than they otherwise would, whenever the topic is being discussed.

Figure 4.1: A graphical model representation of LDA



Source: Blei (2012)

To practically implement the LDA algorithm, researchers need to specify k, the number of topics they wish the algorithm to find. Selecting too few topics can produce results that are too broad, while selecting too many can lead to too detailed and redundant topics. Once k is specified, the algorithm gives the probabilities of words being used in a topic and provides an accounting of the distribution of those topics across the texts. In this paper, we use the Python tutorial code on topic modeling provided by Hansen, McMahon and Prat (2014). As for parameter settings, we follow similar specification used by Hansen and McMahon (2016). Specifically, the model is estimated at the sentence level. By using a collapsed Gibb sampling of Griffiths and Steyvers (2004), we get topic allocation for every iteration of the chain and we draw 20 samples from points in the chain that are thinned by setting an interval of 50. The final topic allocation is given by taking the average of the best-performing 20 samples.

4.1.3 Measuring tone using Dictionary-based method

Generally, dictionary technique is used to measure the intensity of word usage which could capture the content of a corpus. It usually consists of two steps (Bholat et al, 2015); first by defining a list of key words to capture the content of interest and then by representing each document in terms of the (normalized) frequency of words in the dictionary. A prime example of this technique in economics literature is by Baker, Bloom and Davis (2016) which measure economic policy uncertainty from newspapers in the US. Using Boolean search, they count the frequency of news articles containing words related to economic policy and uncertainty. After scaling by total number of articles, Baker, Bloom and Davis construct a proxy for policy uncertainty.

The dictionary techniques have also been extensively applied in the finance and accounting literature. Most prominent examples are Tetlock (2007) and Loughran and McDonald (2011). These papers use dictionary method by classifying positive and negative word to measure tone and sentiment in company report and corporate press release A commonly used source for word classification is the Harvard Psycho-sociological Dictionary, so-called the Harvard-IV-4 list.[12] Then researchers work from this list and modify word list that is suitable for specific research field. In our case, we specifically obtain the list of 'directional words (positive and negative words) used in Hansen and McMahon (2016). The original source of this contraction/expansion list comes from Apel and Grimaldi (2012) that measured tone of monetary policy position (Hawkish vs Dovish) in the minutes of the monetary policy meetings at the Riksbank.

Having obtained the estimates of topic proportions and positive/negative word list, we proceed to measure tone of the policy statements. Our main objective is to examine the tone of the main topics (i.e. economic growth/inflation/monetary policy outlooks). So, we count the (normalized) frequency of expansion/contraction words in these topics at the sentence level and multiplying by estimated proportion of topics (obtained from LDA above). Specifically, we use a simple formula as follows:

[12] http://www.wjh.harvard.edu/~inquirer/homecat.htm

$$Net\ tone = [Share\ of\ topic\ proportion] * \left[\frac{\#\ of\ positive\ words - \#\ of\ negative\ words}{\#\ of\ positive\ words + \#\ of\ negative\ words}\right]$$

It should be noted that there are various ways of measuring tone which could lead to different results. For example, McMahon and Hansen consider uncertainty (or ambiguous words) in the policy statement by dividing the share of uncertain words into the formula. The idea is that the higher the uncertain words, the lower the precision the policy outlook. Furthermore, other papers especially in finance and accounting literature, use more sophisticated techniques by taking into account of word weighting. The main reason for adopting this sophisticated method is the recognition that raw word counts may not be the appropriate measure of a word's information content (Loughran and McDonald, 2011). To address this concern, term-frequency-inverse document frequency (if.idf) has been used to modify the formula (Bholat et al, 2015).

4.2 Empirical results

As shown in the table below, after preprocessing the policy statements of 12 countries, there are 4,418 number of unique stems with 519,605 stem observations in the 1,677 documents over the past 15 years. To visualize these unique stems, figure below provide a word cloud of top 200 stemmed words in our sample.

Table 4.1: Number and frequency of unique words of 12 central bank over 15 years

		Number of unique stems	Number of total stems	Number of documents
Major (non-IT) central banks	FOMC	788	23,074	127
	ECB	2,329	130,539	172
10 IT central banks	Australia	1,094	26,065	108
	Canada	1,115	23,315	123
	Chile	1,049	22,779	179
	Korea	800	27,878	177
	New zealand	1,266	17,323	105
	Norway	1,328	35,887	113
	Peru	905	29,421	179
	Poland	2,027	72,889	174
	South Africa	2,449	93,640	95
	Thailand	1,042	16,795	125
	All 12 countries	**4,418**	**519,605**	**1,677**

Figure 4.2: Word cloud containing top 200 words used by 12 central banks over 15 years

Thus far, the literature on applying topic modeling to central bank communication focuses on a single major central bank, especially in the US. In this paper, our aim is to perform cross-country analysis so as to gain insight on some stylized facts of the main aspects in monetary policy communication. However, estimating topic model for each individual country would result in different topic categories, making it difficult for cross-country comparison. So, our analysis begins with pooling all monetary policy statements of 12 central banks to simultaneously estimate topic categories and to get an overview thematic structure of monetary policy communication. Then we will focus specifically on 10 inflation-targeting central banks (IT10) in order to understand the evolution of topics over time. More in-depth analysis at country level is performed in the case of FED, ECB and the Bank of Thailand (BOT).

As discussed above, one of the key inputs in LDA is to determine the number of topics to be estimated. Following Hansen and McMahon (2016), we use 15-topic model which should be adequate for capturing various issues covered in a short policy statement. Note that we will later show a method for meaningful grouping of related (and probably redundant) topics by using a hierarchical clustering analysis.

4.2.1 Pooling estimation

Table 4.2 shows the top-ten highest probability of words in each topic from our 15-topic model by pooling policy statements of 12 countries. In fact, all words have been assigned probability in all topics with different probability. This is because LDA is a mixed membership model in which words are assigned probabilities and related to multiple topics. And the weights are estimates of $\hat{\beta}_{k,j}$ where k = 0,2,…,14 and represent the probability that word *j* characterizes topic *k*. For each topic, this top-ten key stems best distinguish the topic from all others. As shown in Table 1, it is not difficult for a researcher to determine *topic assignment* based on probability distribution of words within a given topic. It should be emphasized that the LDA algorithm has no information about these topic assignments. The 'interpretable' topics arise mainly from computing the hidden structure in the figure 4.1

above (showing the graphical representation of LDA model). This is one of the main advantages of the LDA model which has the property that inferred hidden structure is able to capture the thematic structure in a collection of the documents and annotates each topic in these documents.

Table 4.2: Distribution of top-ten (stemmed) keywords from 15-topic model

Topic assignment		1	2	3	4	5	6	7	8	9	10
Inflation	topic0	remain	pressur	low	recent	moder	data	continu	growth	inflationari	confirm
		0.083	0.06	0.057	0.054	0.048	0.033	0.032	0.031	0.027	0.024
Labor market	topic1	increas	wage	product	growth	labour	employ	cost	result	adjust	unemploy
		0.091	0.041	0.037	0.034	0.033	0.031	0.028	0.023	0.021	0.02
Domestic demand	topic2	growth	demand	domest	sector	household	continu	credit	invest	loan	privat
		0.088	0.065	0.056	0.048	0.04	0.039	0.036	0.034	0.032	0.028
Economic condition	topic3	econom	committe	condit	inform	support	time	improv	current	assess	gradual
		0.091	0.055	0.046	0.032	0.031	0.03	0.028	0.027	0.023	0.021
Oil prices	topic4	price	increas	oil	consum	rise	food	commod	mainli	energi	declin
		0.274	0.076	0.046	0.039	0.036	0.03	0.025	0.023	0.02	0.018
Risk outlook	topic5	risk	outlook	factor	current	effect	develop	balanc	account	chang	impact
		0.097	0.071	0.038	0.032	0.03	0.029	0.028	0.028	0.027	0.024
Financial market	topic6	market	financi	exchang	intern	us	includ	dollar	rand	volatil	domest
		0.128	0.087	0.042	0.036	0.034	0.025	0.018	0.017	0.017	0.017
Euro area	topic7	euro	area	fiscal	govern	countri	public	financ	need	measur	reform
		0.085	0.084	0.036	0.024	0.024	0.023	0.02	0.019	0.017	0.016
Inflation	topic8	inflat	expect	target	rang	remain	measur	cpi	core	within	end
		0.338	0.106	0.076	0.041	0.037	0.029	0.025	0.024	0.021	0.021
Output growth	topic9	declin	indic	trend	output	growth	show	posit	real	although	improv
		0.059	0.045	0.042	0.032	0.032	0.024	0.022	0.021	0.02	0.019
Global economy	topic10	economi	econom	growth	global	activ	recoveri	export	emerg	world	state
		0.123	0.085	0.081	0.059	0.036	0.034	0.021	0.02	0.018	0.017
Monetary policy	topic11	rate	interest	annual	basi	point	decid	unchang	deposit	key	today
		0.319	0.087	0.068	0.041	0.039	0.038	0.033	0.025	0.024	0.017
Inflation	topic12	term	medium	remain	stabil	develop	expect	continu	close	council	time
		0.109	0.055	0.053	0.053	0.052	0.049	0.044	0.035	0.033	0.028
Monetary policy	topic13	polici	monetari	bank	board	maintain	meet	central	oper	stanc	overnight
		0.168	0.159	0.098	0.043	0.039	0.028	0.026	0.023	0.015	0.012
Growth projection	topic14	project	lower	previou	last	forecast	period	higher	compar	gdp	averag
		0.062	0.053	0.04	0.037	0.037	0.037	0.036	0.034	0.032	0.029

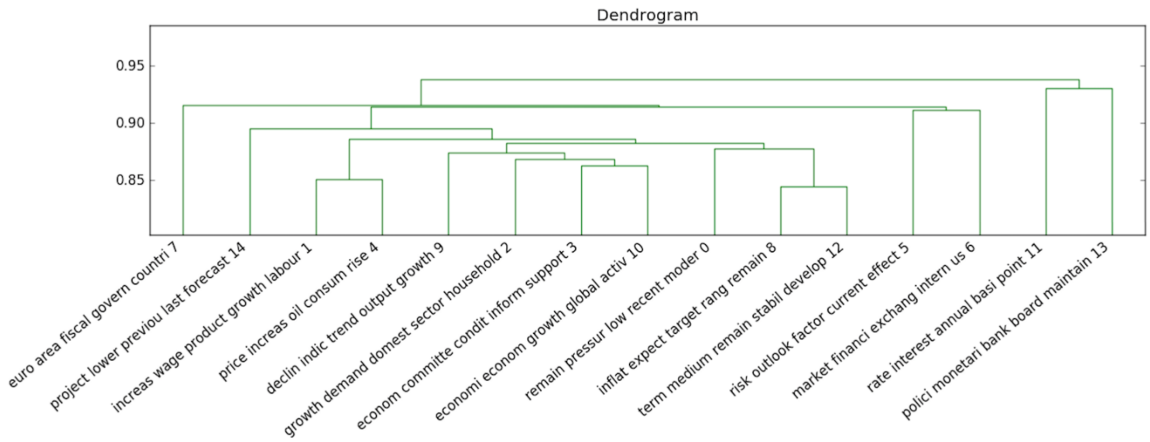
Stylized fact 5: A thematic view reveals that monetary policy is a complicated decision-making

How does monetary policy committee explain its policy decision? A thematic view

By estimating topic proportion as well as word distributions, we then use hierarchical clustering analysis to construct the structure of central bank rhetoric. The main purpose is to see how these 15 topics are related to each other. Note that hierarchical clustering analysis has actively been applied to political science, see Grimmer (2010) and Quinn et al. (2010), for example. How the monetary policy committee explains its decision in the press statement is shown in the form of a tree-based hierarchical diagram or 'dendrogram' in the figure below. The vertical axis represents Euclidean distance. The lower the height at which any two topics are connected, the more similar are their word use patterns. Reading the figure from the bottom-up shows information about which clusters are merged first at the lowest height. We can see that topic 8 and topic 9 are basically in the same category about inflation that are merged first due to most similarity in the word usage. Then topic 0 belongs to the same inflation cluster. Topics 1 and 4 are another clustering set about energy prices and wage growth in the labor market. Then, topics 9,2,3, and 10 are essentially about economic growth condition.

This figure illustrates graphically an important feature of the policy statement. That is, monetary policy is a complicated decision-making and the committee needs to balance a number of economic issues when deciding appropriate policy rate for the economy. When viewed this dendrogram from the top-down, we can see that topics 11 and 13 are clearly about monetary policy or policy interest rate decision. The committee then look at risk outlook and financial markets (T5 and T6) and assessing its economic projection (T14) relative to three main clusters on wage and prices (T1 and T4); the real economy (T9,T2,T3,T10); and inflation (T8,12).

Figure 4.3: Tree-based hierarchical taxonomy of the policy statements



Based on the dendrogram, we can re-clustering these 15 topics by grouping the same cluster into 6 main topics for further analysis. Specifically, these are: Monetary policy (T11,T13); Financial market (T5,T6); Inflation (T0,T8,T12); Growth (T9,T2,T3,T10); Price (T1,T4); Euro Area (T7). And the results are shown in the form of word clouds in the figure below. On the examination of these word clouds, we can see that these 6 topics can be labeled easily. So, the dendrogram not only allows us to understand the semantic inter-relationship among various topics, but it also help us to decide how to simplify many topics into a new set of manageable and meaningful categories.

Figure 4.4: Word clouds of re-grouping 15 topics into 6 main topics

[Monetary policy: T11, T13]

[Financial market: T5, T6]

[Inflation: T0,T8,T12]

[Economic growth: T9,T2,T3,T10]

[Prices: T1,T4]

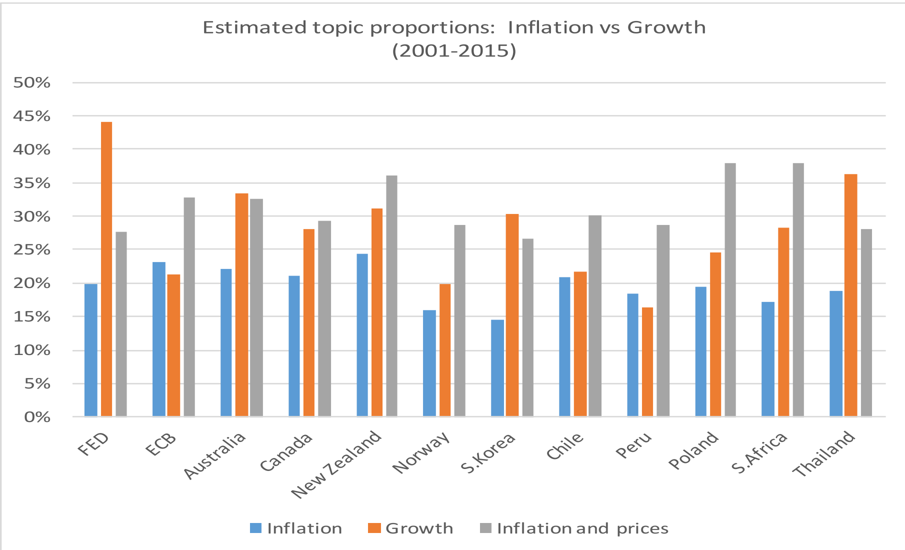
[Euro area: T7]

4.2.2 How do inflation-targeting countries talk?

Stylized fact 6: Most IT countries discussed more about inflation-related topics in the policy statements

We are now able to compare among ten IT countries with two major central banks by first asking whether inflation targeting countries communicate more about inflation, relative to other topics especially the economic growth topic in their policy statement. Figure 4.5 shows the average of estimated topic proportions on inflation, prices and growth over the past 15 years. As can be seen in the figure, most countries in our sample talk much more about economic growth, with the FED having highest proportion of around 45% in the policy statement while ECB having quite balanced statement between growth and inflation. When topic on prices has been included to inflation topic, however, the picture changes substantially. Almost all IT countries discussed more about inflation-related topics in the policy statements, except Thailand and South Korea (one can include Australia but the difference is very small).

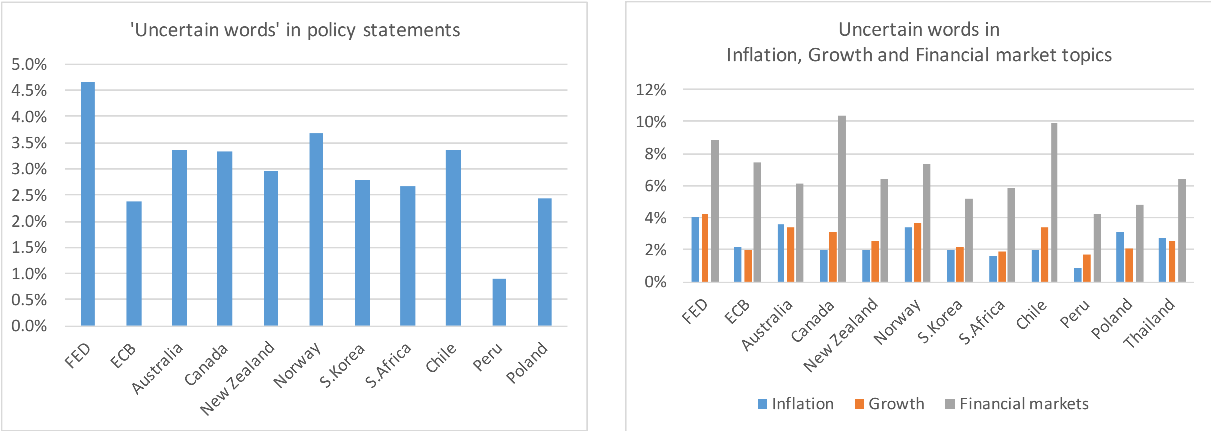
Figure 4.5: Average of estimated topic proportions during 2001-2015



Stylized fact 7: Use of uncertainty words in the policy statements are expressed much more frequent in financial market topic, relative to inflation and growth topics.

In addition, using the ambiguity word list from Hansen and McMahon (2016) [the word list was originally developed by Loughran and McDonald, 2011], we use automated dictionary method by counting the number of uncertain words at the sentence level of every policy statements and normalized by total number of stems in each sentence. The total sentence of the policy statements of 12 central banks over the past 15 year is 44,823. We find that the FOMC statement of the FED, relative to other eleven countries contains highest uncertain words (4.6%). To put this into perspective, Loughran and McDonald (2011) reported that the US corporate annual report (so-called 10-K report) generally contains uncertain words averaging of around 1.2% (using data between 1994-2008). When looking specifically at the topic level, the use of uncertainty words in the policy statements are expressed much more frequent in financial market topic, relative to inflation and growth topics. See Figure 4.6 below.

Figure 4.6: Use of uncertain words in the policy statements



In order to investigate more about topic evolution in the monetary policy communication, we re-run the 15 topic model by excluding FED and ECB and selecting only 10 IT economies. Using the same methodology as discussed above (i.e. re-clustering by dendrogram), the simplified thematic structure of the monetary policy statements is shown in Figure X. The figure reveals clearly that monetary policy (MP) in these countries need to find a striking balance between domestic economy (i.e. growth, inflation and prices, see the blue lines) and external environments (namely exchange rate, risk, global economy and global financial markets, see the green lines). Figure 4.7 below gives details of the word distribution in these 8 topics.

Figure 4.7: Dendrogram of the policy statements in 10 inflation-targeting countries

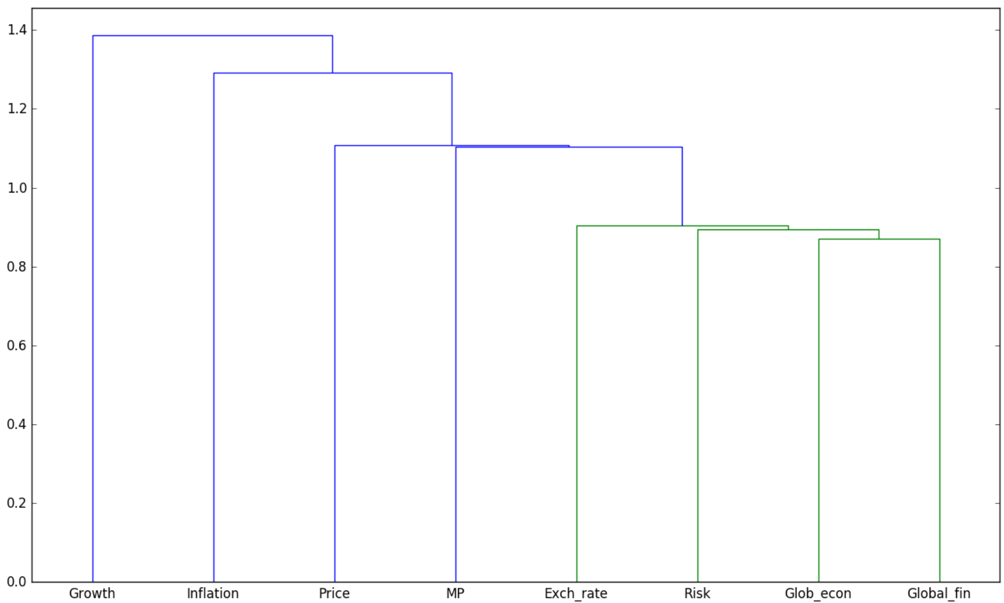


Figure 4.8: Word clouds of 8 main topics in the IT10

Stylized fact 8: Since 2008, discussions of global topics become more prominent in all IT countries

Tables below illustrate how the topic discussions in IT10 evolve between pre and post global financial crisis of 2008 (GFC). We find that role of the 'global factors' becomes more prominent in all IT countries both advanced and emerging economies, especially topics about global economy and global financial market. The proportion of discussion on domestic economy in Norway declined sharply by 13% and the share of monetary policy topic increased from 18% to 30% (not shown in the table). South Korea has also seen significant reduction in topic proportion of domestic economy and discussed more actively about the external global environments, especially the topics about global financial market condition. In the case of Thailand, the discussion of global economy increased its proportion significantly to 16.1% in the post GFC from 9.3% during the pre-GFC period.

Table 4.3: Topic proportions of domestic economy VS 'global factors'

	Domestic economy			External environment		
	2001-2007	2008-2015	Difference	2001-2007	2008-2015	Difference
Australia	57%	55%	-2%	31%	33%	2%
Canada	53%	52%	0%	26%	32%	7%
New Zealand	60%	58%	-2%	31%	34%	3%
Norway	58%	45%	-13%	24%	25%	1%
S.Korea	67%	52%	-15%	17%	35%	19%
Chile	53%	48%	-5%	20%	27%	7%
Peru	52%	49%	-2%	14%	15%	1%
Poland	67%	59%	-7%	20%	25%	4%
S. Africa	69%	65%	-4%	23%	29%	6%
Thailand	57%	51%	-7%	31%	37%	6%

Table 4.4: Topic proportions of global factors

	Global economy		Global financial market		Exchange rate		Risk outlook	
	2001-2007	2008-2015	2001-2007	2008-2015	2001-2007	2008-2015	2001-2007	2008-2015
Australia	8.6%	8.4%	5.8%	8.2%	9.9%	11.3%	7.0%	4.9%
Canada	10.3%	14.6%	3.1%	5.7%	4.3%	5.3%	7.9%	6.8%
New Zealand	9.6%	11.0%	4.5%	6.3%	9.1%	9.9%	8.0%	6.6%
Norway	6.1%	9.9%	4.0%	4.9%	7.8%	5.5%	6.0%	4.9%
S.Korea	2.7%	5.9%	9.6%	19.3%	1.9%	5.6%	2.6%	4.5%
Chile	4.7%	6.9%	5.0%	7.8%	4.8%	4.2%	5.6%	8.0%
Peru	1.3%	2.7%	3.1%	4.4%	6.1%	3.2%	3.2%	4.8%
Poland	2.7%	5.6%	5.7%	10.5%	3.7%	3.2%	8.4%	5.5%
S. Africa	4.0%	7.2%	3.9%	5.4%	8.7%	8.3%	6.6%	7.9%
Thailand	9.3%	16.1%	5.8%	7.8%	6.0%	4.5%	10.1%	9.0%

5. What determines topic proportions and tones?

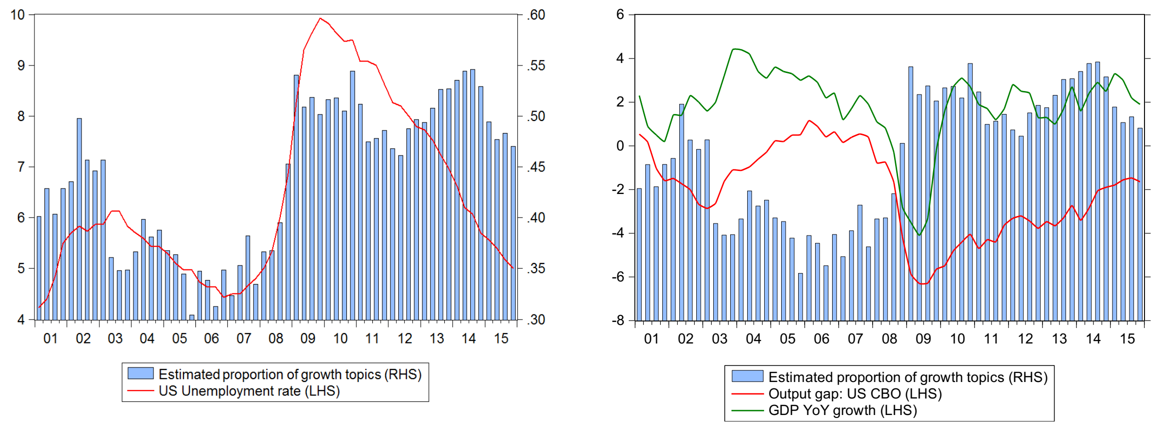
Stylized fact 9: There is a diversity in the degree of informativeness when communicating about economic situations

In this section, we investigate how our estimates of topic proportions and tone are related to some key macroeconomic variables, i.e. economic growth, inflation and policy interest rate. Our objective is to examine the informativeness of both topics and tones in the policy statements. We focus specifically at three central banks, i.e. FED, ECB and Thailand, respectively.

5.1 US Federal Reserve

We begin with the analysis of growth topics in the FOMC statement of the FED. Figure 5.1 shows the evolution of estimated topic proportion, compared with actual GDP year-on-year growth, output gap and unemployment rate. As clearly seen in the figure, the proportion of growth topics increased sharply since 2008, averaging around 50% of the total proportion. We find that our estimated growth topic proportion is related negatively to GDP growth and positively to unemployment rate. An eye-balling test indicate a contemporaneous correlation between economic data and the amount of language used to discussed the topic. Our cross correlation analysis, however, suggests that the topic proportion on growth seems to lag the actual statistics, implying that the proportion of growth topics in the FOMC statements relies more on incoming data, rather than reflecting economic growth outlook.

Figure 5.1: Proportion of growth topics VS GDP growth, Unemployment and Output Gap

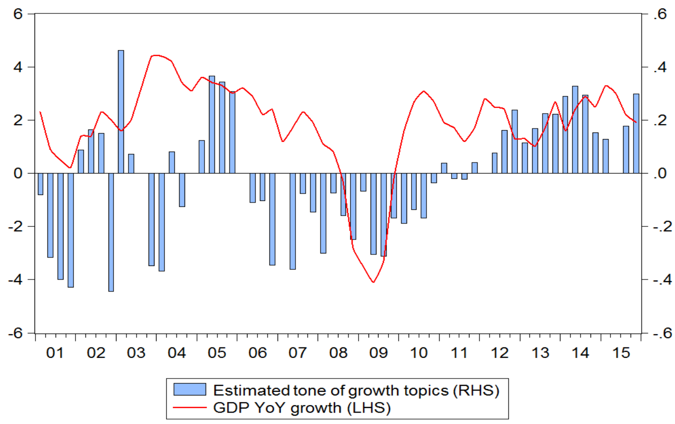


Sample: 2001Q1 2015Q4
Included observations: 60
Correlations are asymptotically consistent approximations

Estimated proportion of growth topics,GDPYoY(-i)	Estimated proportion of growth topics,GDPYoY(+i)	i	lag	lead
		0	-0.32	-0.32
		1	-0.38	-0.21
		2	-0.43	-0.08
		3	-0.48	0.08
		4	-0.51	0.21
		5	-0.51	0.29
		6	-0.51	0.32
		7	-0.47	0.32
		8	-0.42	0.30

The net tone of growth topics is shown in Figure 5.2 below. While our estimate of tone on growth topics correlated more contemporaneously with GDP growth, the lead-lag structure indicates that the net tone slightly leads actual GDP growth. As for the inflation topics, we also find that the topic correlates quite closely with actual movements of inflation (PCE YoY growth). Interestingly, as shown in Figure 5.3, the proportion of the inflation topic is more informative about inflation outlook than the net tone.

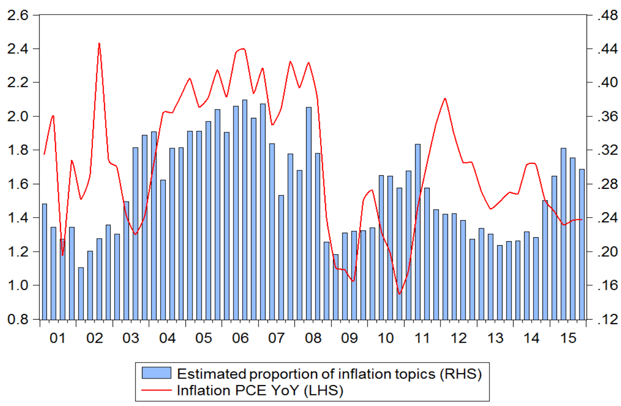
Figure 5.2: Net tone of growth topics VS GDP YoY growth



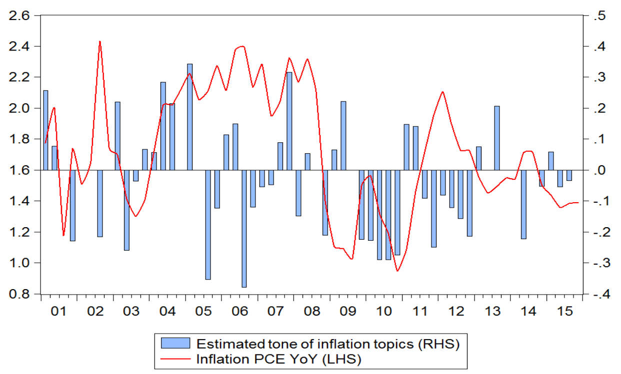
Sample: 2001Q1 2015Q4
Included observations: 60
Correlations are asymptotically consistent approximations

Estimated tone of growth topics,GDPYoY(-i)	Estimated tone of growth topics,GDPYoY(+i)	i	lag	lead
		0	0.31	0.31
		1	0.30	0.31
		2	0.31	0.29
		3	0.26	0.28
		4	0.24	0.23
		5	0.18	0.21
		6	0.15	0.22
		7	0.06	0.19
		8	0.06	0.16

Figure 5.3: Proportion and Tone of inflation topics VS Actual inflation (PCE)



Estimated proportion of inflation topics, Inflation YoY (-i)	Estimated proportion of inflation topics, Inflation YoY (+i...	i	lag	lead
		0	0.48	0.48
		1	0.38	0.49
		2	0.33	0.50
		3	0.29	0.54
		4	0.22	0.52
		5	0.18	0.56
		6	0.13	0.54
		7	0.03	0.43
		8	0.04	0.26

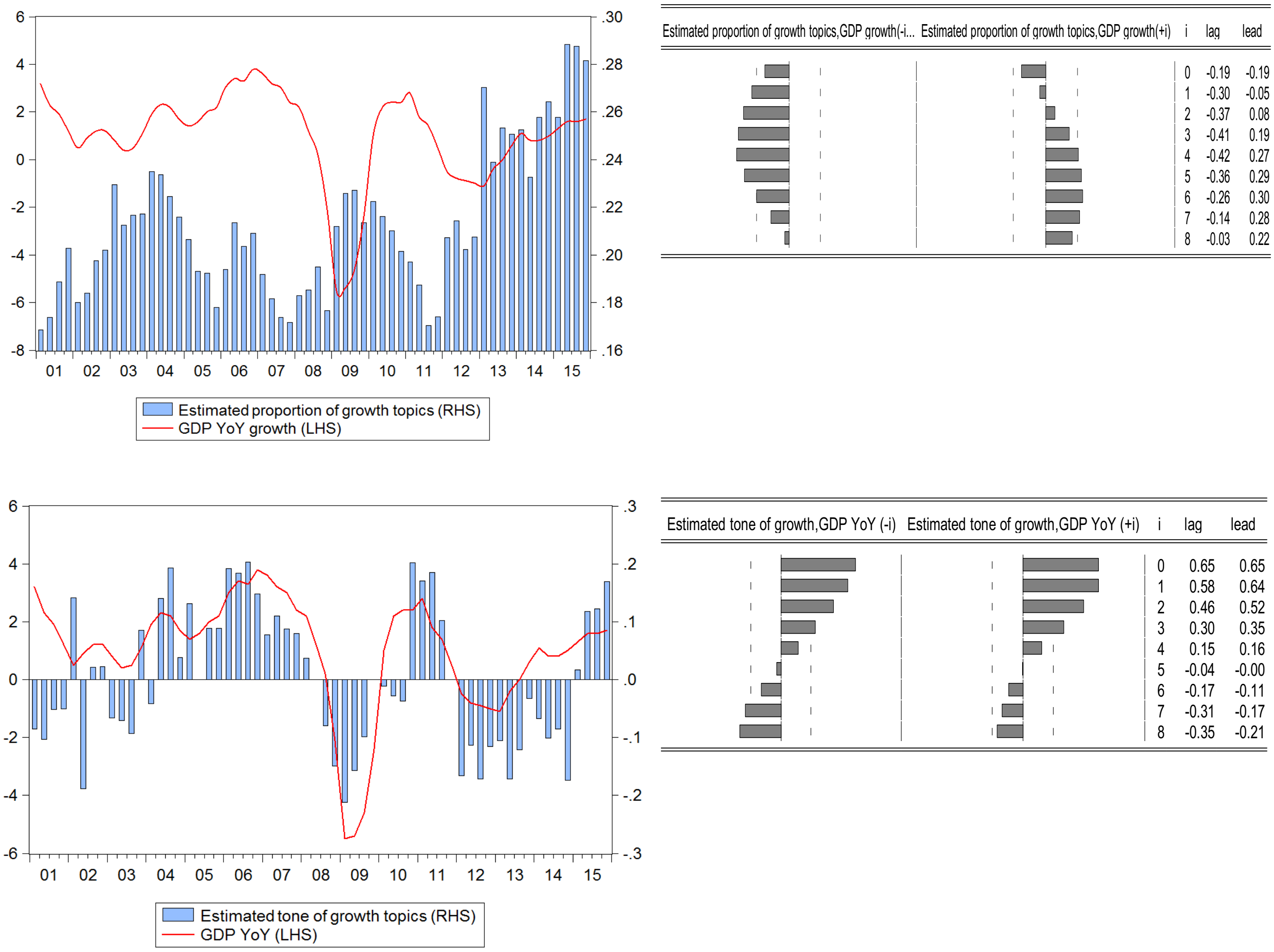


Estimated tone of inflation topics, Inflation YoY (-i...	Estimated tone of inflation topics, Inflation YoY (+i...	i	lag	lead
		0	0.06	0.06
		1	-0.01	0.17
		2	0.01	0.21
		3	0.05	0.25
		4	0.18	0.11
		5	0.23	0.03
		6	0.12	0.03
		7	0.14	-0.08
		8	0.02	0.05

5.2 European Central Bank

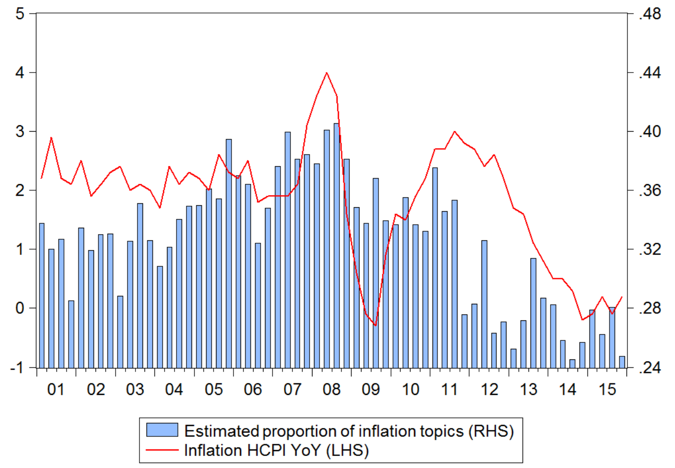
Figure 5.4 shows the proportion and tone of growth topics in the ECB's policy statement, compared with actual GDP growth in the Euro area. As shown in the figure, lower GDP growth leads to more discussion about it and the proportion devoted to growth topics increased significantly since 2012. Our estimates of net tone on growth topics highly correlate with actual GDP growth with slightly leading the actual figures. The correlation is as high as 65%.

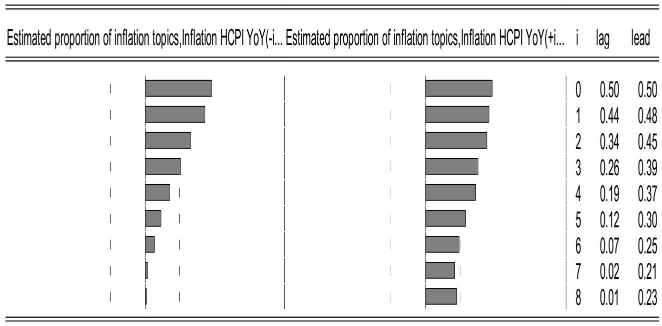
Figure 5.4: Proportion and tone of growth topics vs GDP growth in the Euro area



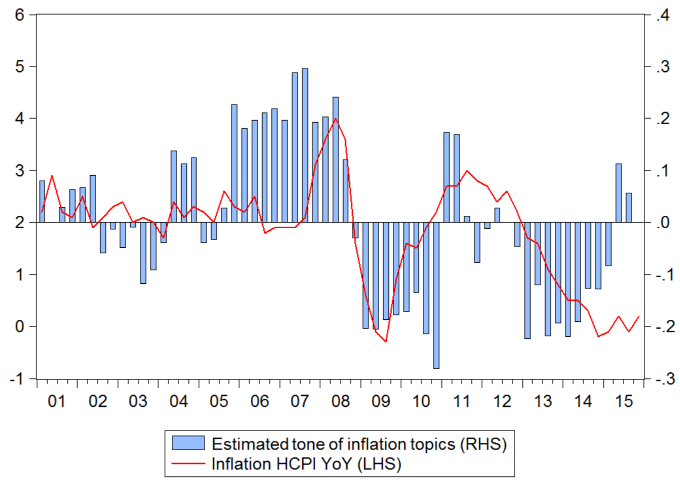
Regarding the topics of inflation in the ECB's policy statement, Figure 5.5 shows that our estimated proportion devoted to inflation topics has declined significantly from the 40% during 2008 to only around 24-28 percent during 2011-2015. In addition, the proportion of inflation declines as actual rate of inflation falls. Looking at the cross-correlation relationships, we find that both estimated net tones of inflation and growth in the ECB policy statements are more informative when compared with the FOMC statements.

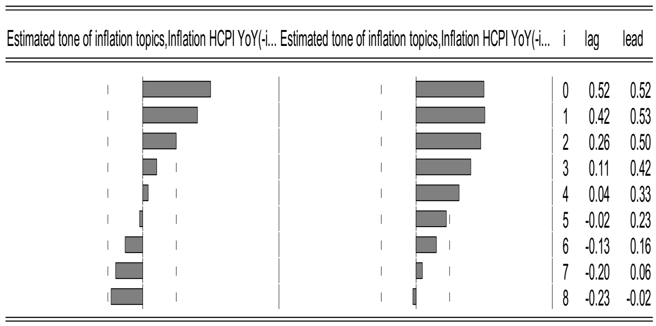
Figure 5.5: Proportion and tone of inflation topics vs actual inflation in the Euro area





Estimated proportion of inflation topics,Inflation HCPI YoY(-i...	Estimated proportion of inflation topics,Inflation HCPI YoY(+i...	i	lag	lead
		0	0.50	0.50
		1	0.44	0.48
		2	0.34	0.45
		3	0.26	0.39
		4	0.19	0.37
		5	0.12	0.30
		6	0.07	0.25
		7	0.02	0.21
		8	0.01	0.23





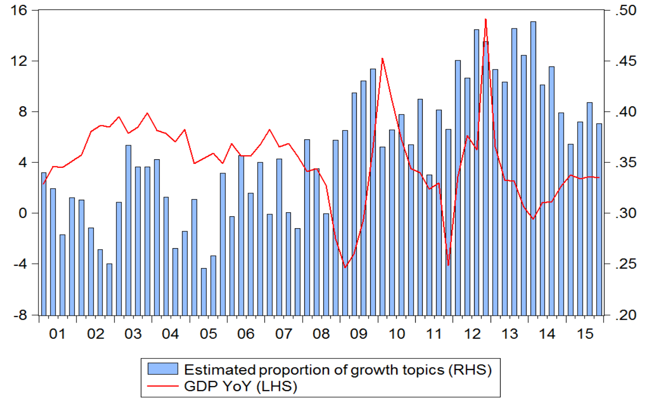
Estimated tone of inflation topics,Inflation HCPI YoY(-i...	Estimated tone of inflation topics,Inflation HCPI YoY(-i...	i	lag	lead
		0	0.52	0.52
		1	0.42	0.53
		2	0.26	0.50
		3	0.11	0.42
		4	0.04	0.33
		5	-0.02	0.23
		6	-0.13	0.16
		7	-0.20	0.06
		8	-0.23	-0.02

5.3 Bank of Thailand

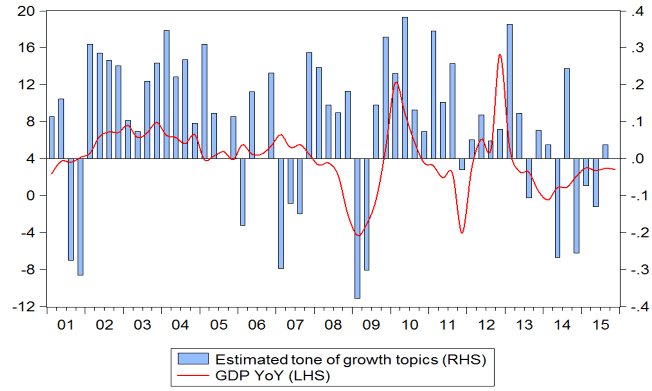
Figure 5.6 show the variation of growth topic proportions comparing with actual GDP growth. The estimated proportion of growth-related topics has generally been increasing over the past 15 years. We also find that the Monetary Policy Committee has tendency to talk more about growth in the statement when the observed GDP growth declines. In addition, the net tone on growth in the policy statements appears to be backward-looking as it is correlated more with one-period lagged GDP growth (see the left panel).

Regarding the inflation topics, Figure 5.7 shows that the average proportion had been declined since 2008 but began to rise when disinflation became more prominent during 2014-2015. We find that both proportion and net tone on inflation related topics appears to be forward looking and more informative relative to growth topics.

Figure 5.6: Growth topics in the BOT policy statements

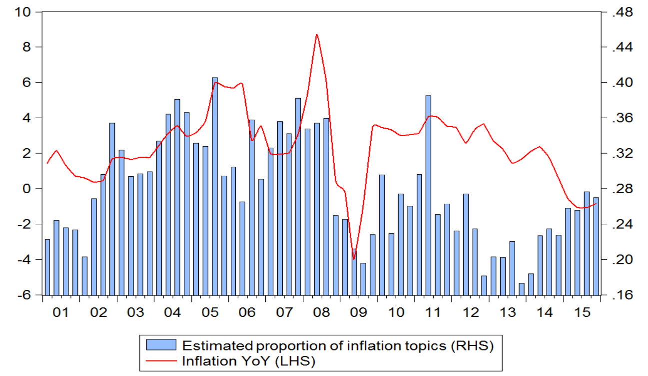


Estimated proportion of growth topics,GDP YoY(-i)	Estimated proportion of growth topics,GDP YoY(-i)	i	lag	lead
		0	-0.25	-0.25
		1	-0.31	-0.15
		2	-0.30	-0.10
		3	-0.28	-0.14
		4	-0.23	-0.22
		5	-0.17	-0.25
		6	-0.20	-0.34
		7	-0.17	-0.25
		8	-0.12	-0.28

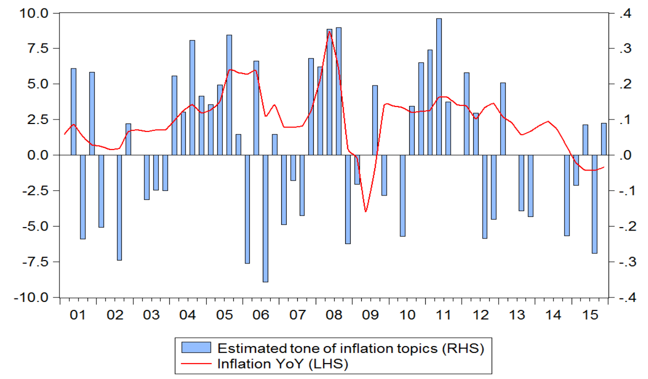


Estimated tone of growth topics,GDP YoY(-i)	Estimated tone of growth topics,GDP YoY(+i)	i	lag	lead
		0	0.36	0.36
		1	0.44	0.27
		2	0.20	0.12
		3	0.09	-0.16
		4	0.09	-0.24
		5	-0.04	-0.06
		6	-0.10	0.01
		7	-0.02	0.21
		8	-0.20	0.20

Figure 5.7: Inflation topics in the BOT policy statements



Estimated proportion of inflation topics,Inflation YoY(-i...	Estimated proportion of inflation topics,Inflation YoY(+i...	i	lag	lead
		0	0.43	0.43
		1	0.32	0.44
		2	0.12	0.42
		3	-0.02	0.37
		4	-0.04	0.28
		5	-0.11	0.32
		6	-0.04	0.29
		7	-0.11	0.27
		8	-0.15	0.21

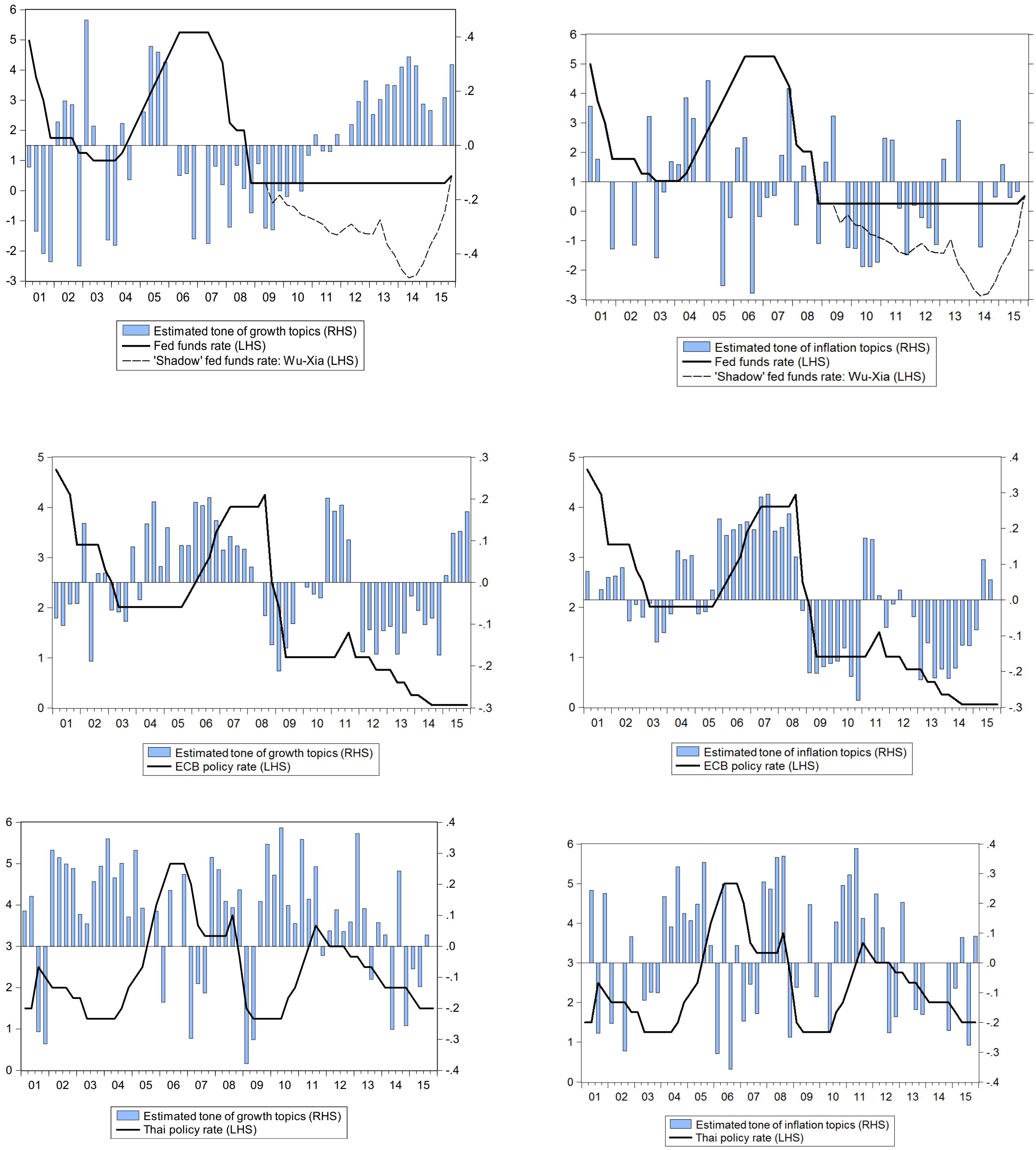


Estimated tone of inflation topics,Inflation YoY(-i...	Estimated tone of inflation topics,Inflation YoY(+i...	i	lag	lead
		0	0.40	0.40
		1	0.16	0.36
		2	0.04	0.35
		3	-0.02	0.10
		4	0.04	-0.05
		5	-0.06	-0.01
		6	-0.13	0.05
		7	-0.21	0.11
		8	-0.24	0.11

5.4. Linking interest rate decision to policy statements

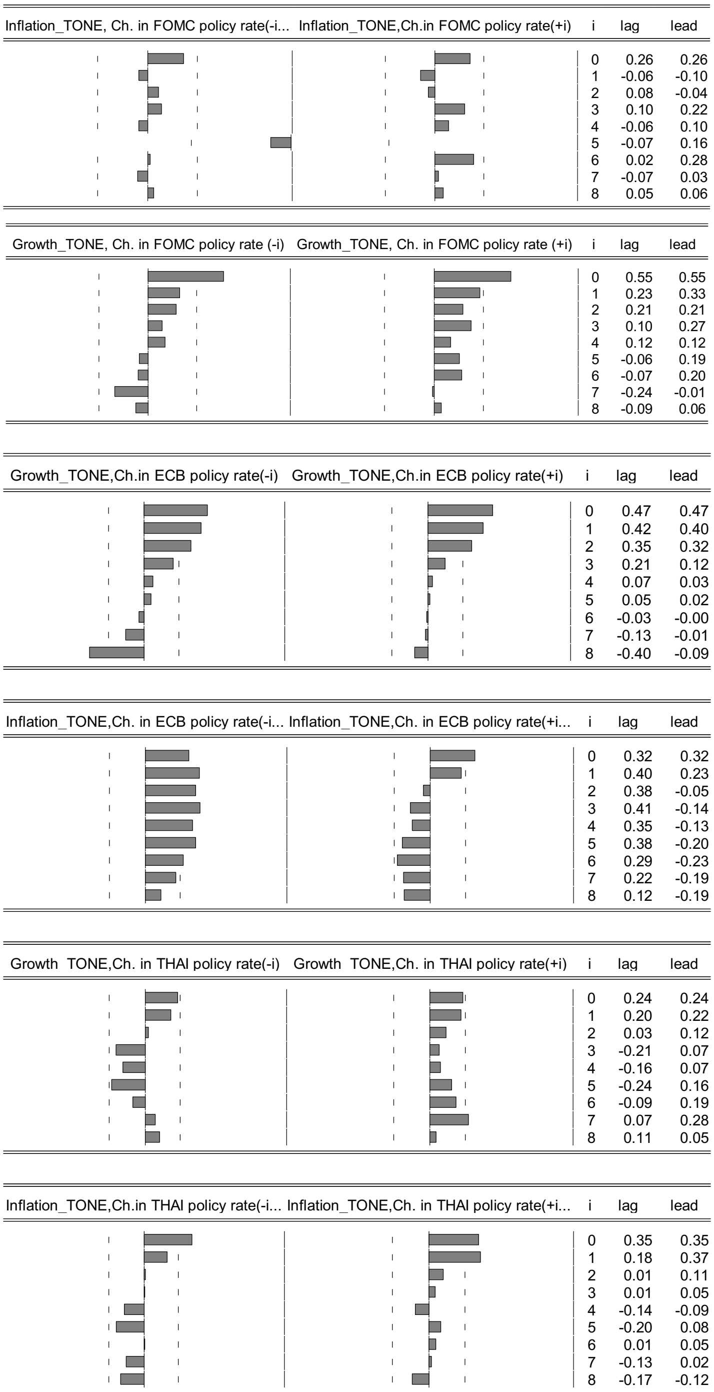
Figure 5.8 shows the relationship between policy interest rate decision and the net tone of inflation and growth topics of the FED, ECB and BOT. As shown in the figure, the ECB policy rate decision is highly consistent with both growth and inflation tones.

Figure 5.8: Policy interest rates vs Growth and Inflation tones



Since it is hard to determine objectively how our estimated net tones and actual interest rate decision are related, we conduct a cross-correlation between change in interest rate (Δi_t) and tone of inflation and growth topics, see Figure 5.9. Regarding the FED, the fed funds rate is significantly related to net tone in economic growth topics but not inflation tone. This is just the opposite to the Bank of Thailand as we find that Thai policy rate is related more to inflation tone. In the case of the ECB, its policy rate is correlated significantly both to growth and inflation tones.

Figure 5.9: Cross correlation between Tone and Policy rate



FOMC

Inflation_TONE, Ch. in FOMC policy rate(-i... Inflation_TONE,Ch.in FOMC policy rate(+i)

i	lag	lead
0	0.26	0.26
1	-0.06	-0.10
2	0.08	-0.04
3	0.10	0.22
4	-0.06	0.10
5	-0.07	0.16
6	0.02	0.28
7	-0.07	0.03
8	0.05	0.06

Growth_TONE, Ch. in FOMC policy rate (-i) Growth_TONE, Ch. in FOMC policy rate (+i)

i	lag	lead
0	0.55	0.55
1	0.23	0.33
2	0.21	0.21
3	0.10	0.27
4	0.12	0.12
5	-0.06	0.19
6	-0.07	0.20
7	-0.24	-0.01
8	-0.09	0.06

ECB

Growth_TONE,Ch.in ECB policy rate(-i) Growth_TONE,Ch.in ECB policy rate(+i)

i	lag	lead
0	0.47	0.47
1	0.42	0.40
2	0.35	0.32
3	0.21	0.12
4	0.07	0.03
5	0.05	0.02
6	-0.03	-0.00
7	-0.13	-0.01
8	-0.40	-0.09

Inflation_TONE,Ch. in ECB policy rate(-i... Inflation_TONE,Ch. in ECB policy rate(+i...

i	lag	lead
0	0.32	0.32
1	0.40	0.23
2	0.38	-0.05
3	0.41	-0.14
4	0.35	-0.13
5	0.38	-0.20
6	0.29	-0.23
7	0.22	-0.19
8	0.12	-0.19

Bank of Thailand

Growth TONE,Ch. in THAI policy rate(-i) Growth TONE,Ch. in THAI policy rate(+i)

i	lag	lead
0	0.24	0.24
1	0.20	0.22
2	0.03	0.12
3	-0.21	0.07
4	-0.16	0.07
5	-0.24	0.16
6	-0.09	0.19
7	0.07	0.28
8	0.11	0.05

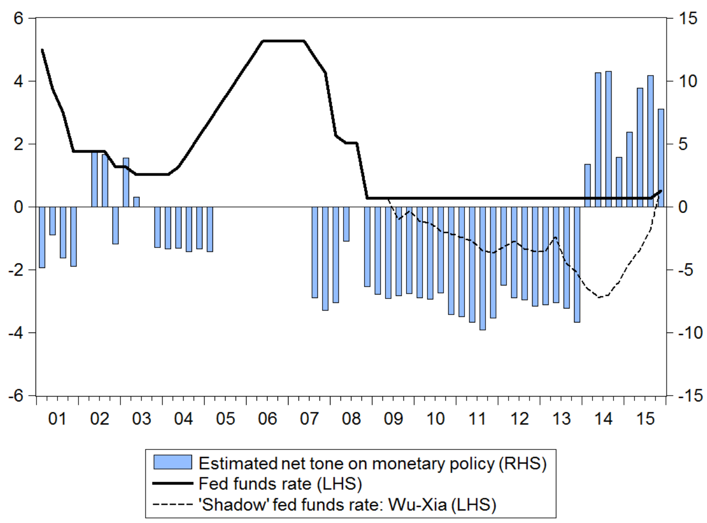
Inflation_TONE,Ch.in THAI policy rate(-i... Inflation_TONE,Ch. in THAI policy rate(+i...

i	lag	lead
0	0.35	0.35
1	0.18	0.37
2	0.01	0.11
3	0.01	0.05
4	-0.14	-0.09
5	-0.20	0.08
6	0.01	0.05
7	-0.13	0.02
8	-0.17	-0.12

Lastly, we look specifically at US monetary policy. As shown in Figure 5.8, fed funds rate in the US had reached zero-lower bound for 7 years since December 2008 (until 15 December 2015). To manage public expectations about the future path of interest rates, the FOMC had to resort to unconventional policy by not only implementing the large-scale asset purchases (LSAP) but also giving the forward guidance on monetary policy through the use of language in the FOMC statement. In order to measure the stance of monetary policy stance during this zero-lower bound, we use a 'shadow' fed funds rate, computed by Wu and Xia (2014).

As shown the Figure 5.10, our estimated net tone on monetary policy topics in the US shows that the FOMC has increasingly been actively used language in the statement to signal monetary policy outlook after its policy rate reaching the zero-lower bound. In addition, when the FED changed its monetary policy tone in early 2014, the shadow rate appears to move accordingly. This indicates that the LDA algorithm is able to objectively capture the language tone in the policy statements. And this can be used for further investigation, for example the macroeconomic effectiveness of forward guidance.

Figure 5.10: Estimated tone on US monetary policy



6. Conclusions

This paper employed a sophisticated textual analysis to measure directly and objectively the written languages of monetary policy statements. We demonstrated that tools from computational linguistics offered a new lens for uncovering hidden complexity in monetary policy making and communication. The paper contributed to existing literature in two main directions. First, we combine multiple tools from computational linguistics to examine various aspects of central bank communication, namely (i) the overall clarity via readability measures (ii) specific economic issues emphasized in the policy statements using topic modeling and (iii) assessment of economic outlook by tone analysis. Second, instead of focusing on single central bank, we look at cross-country level to obtain stylized facts about overall evolution of central bank communication during the past 15 years. Our findings can be highlighted as follows:

- Across 22 central banks, information disclosure in the monetary policy statements has increased steadily and the texts have become clearer. This is a welcoming trend. Nevertheless, the documents are still far from being light readings. In addition, there is a substantial degree of cross-sectional variation in terms of readability. In particular, documents from three major central banks, namely FED, ECB and BOJ, are quite complex. On the other hands, Bank of Norway is the best for its writing clarity.
- An attribution analysis shows that syntactic complexity carries the largest weight on the overall complexity score, followed by components that represent vocabulary difficulty, namely academic vocabulary, word unfamiliarity, and concreteness.
- Readability tends to fall when central banks lower their policy interest rates. Statistically speaking, for every 25bps cut, a statement will contain around 5 more words per statement on average.
- As for inflation-targeting countries, the policy statements in this group generally contained more discussions about inflation-related topics, relative to growth topics. This is not surprising. Nevertheless, we find that discussions of global topics become more prominent in aftermath of 2008 financial crisis. Key topics on the global developments include global economy and global financial markets. For example, in the case of Thailand, the discussion of global economy increased its proportion significantly to 16.1% in the post 2008 from 9.3% during the pre-2008 period.
- Looking at topic evolution of three central banks, i.e. FED, ECB and BOT, our estimated topic proportions are generally consistent with actual economic data. In particular, we find that the proportion on growth topics is negatively related to GDP growth while the proportion on inflation topics is positively related to actual change in inflation.
- When it comes to giving assessment on the economy, the FOMC policy statements are only informative on growth while the ECB statements are highly consistent and informative both on inflation and growth outlook. Net tone on growth in the Thai policy statements appears to be backward-looking as it is correlated more with one-period lagged GDP growth. But the assessment of inflation outlook appears to be forward looking and so more informative.
- Tone matters. Specifically, when the net tone is informative about economic outlook, it is also significantly related to change in policy rate decision.

Our paper suggests a policy implication for designing monetary policy communication. There are two important necessary requirements for effective communication. (i) Being clear. The language used in the policy statements needs to be simple. While there are various dimensions contributing to language complexity, our result indicates that academic or technical words should be avoided. (ii) Being informative. The proportion of concerned topics and tone of economic assessment should be more forward looking.

Balancing between clarity and informativeness, however, will remain a challenging task for central banks. As remarked by Otmar Issing (2014):

"Experience has shown how difficult it is to communicate all information relevant to the decision making process in a way that is not only exhaustive but also clear and comprehensible. Psychological research has pointed to the limits of human information processing skills (Kahneman, 2003). This research has shown, e.g. that the weighing of information greatly depends on its intuitive accessibility. Furthermore, it is generally simplified and categorized before it is collated. Altogether this implies a huge challenge for central bank communication. A too sophisticated approach could easily create confusion rather than providing clarity."

Furthermore, Jeremy Stein (2014), offered view based on the experience at the FOMC:

"As we evaluate our own performance in the communications department, it is probably better for us to focus on how legitimately transparent we have succeeded in being, as opposed to how much or how little our various announcements have moved markets."

References

Apel, M., and M. B. Grimaldi (2012), 'The information content of central bank minutes', Working Paper Series 261, Sveriges Riksbank, April.

Bholat, David (2014), 'Big data and central banks', Mimeo.

Bholat, David, Stephen Hansen, Pedro Santos, and Cheryl Schonhardt-Bailey (2015), 'Text mining for central banks', Centre for Central Banking Studies Handbook No.33, The Bank of England.

Blei, D. M., N. Ng, and M. Jordan (2003), 'Latent Dirichlet Allocation', *Journal of Machine Learning Research*, 3: 993-1022.

Blei, D. M. (2012), 'Probabilistic topic models', *Communications of the ACM*, 55(4): 77-84.

Blinder, Alan S., Michael Ehrmann, Marcel Fratzscher, Jakob de Hann, and David-Jan Jansen (2008), 'Central bank communication and monetary policy: A survey of theory and evidence', *Journal of Economic Literature*, 46(4): 910-945.

Bulíř, Aleš, Martin Čihák, and David-Jan Jansen (2013), 'What drives clarity of central bank communication about inflation?', *Open Economies Review*, 24(1): 125-145.

Dennis, Richard, and John C. Williams (2007), 'Monetary policy, transparency, and credibility: Conference summary', Federal Reserve Bank of San Francisco Economic Letter, 2007-12.

Dincer, N. Nergiz, and Barry Eichengreen (2014), 'Central Bank Transparency and Independence: Updates and New Measures', *International Journal of Central Banking,* 10(1): 189-253.

DuBay, William H (2006), *The Classic Readability Studies*. CA: Impact Information.

Ehrmann, Michael, and Marcel Fratzscher (2005), 'How should central banks communicate?', ECB Working Paper No. 557, November 2005.

Flesch, Rudolf (1948), 'A new readability yardstick', *Journal of Applied Psychology*, 32(3): 221-233.

Fracasso, Andrea, Hans Genberg, and Charles Wyplosz (2003), 'How do central banks write?', *Geneva Reports on the World Economy*, Special Report 2, Centre for Economic Policy Research, Geneva.

Griffiths, T. L., and M. Steyvers (2004), 'Finding scientific topics', *PNAS*, 101: 5228-5235.

Grimmer, Justin and Brandon M. Stewart (2013), 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts', Political Analysis, 21(3), pp.1-31.

Hansen, Stephen, Michael McMahon, and Andrea Prat (2014), 'Transparency and deliberation with the FOMC: A computational linguistic approach', CEP Discussion Paper No.1276, Centre for Economic Performance, LSE.

Hansen, Stephen, and Michael McMahon (2016), 'Shocking language: Understanding the macroeconomic effects of central bank communication', *Journal of International Economics*, 99: 114-133.

Holmes, Douglas R. (2014), *Economy of Words: Communicative Imperatives in Central Banks*. Chicago: The University of Chicago Press.

Hughes, Peter T., and Stefan Kesting (2014), ‘A literature review on central bank communicaiton’ *On the Horizon*, 22(4): 328-340.

Irving Fisher Committee on Central Bank Statistics (2015), ‘Central banks use of and interest in big data’, IFC Report, October 2015.

Jansen, David-Jan (2011). ‘Does the clarity of central bank communication affect volatility in financial markets? Evidence from Humphrey-Hawkins testimonies’, *Contemporary Economic Policy*, 29(4): 494-509.

Jansen, David-Jan, & de Haan, Jakob (2010). An assessment of the consistency of ECB communication using Wordscores. DNB Working Paper, No. 259 (April).

Kedan, Danielle, and Rebecca Stuart (2014), ‘Central bank communications: A comparative study’, *Central Bank of Ireland Quarterly Bulletin*, No.02/April 2014: 89-104.

Kincaid, J. P., R. P. Fishburne, R. L. Rogers, and B. S. Chissom (1975), ‘Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel’, Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis.

Levin, Andrew T. (2014), ‘The design and communication of systematic monetary policy strategies’, *Journal of Economic Dynamics & Control*, 49: 52-69.

Lorge, Irving (1944), ‘Predicting readability’, *Teachers College Record*, 45: 404-419.

Lorge, Irving (1948), ‘The Lorge and Flesch Readability Formulae: A Correction’, *School and Society*, 67: 141-142.

Loughran, T., and B. McDonald (2011), ‘When is a liability not a liability? Textual analysis, dictionaries and 10-Ks’, *Journal of Finance*, 66(1): 198-210.

Luangaram, Pongsak and Yuthana Sethapramote (2016), ‘Central bank communication and monetary policy effectiveness: Evidence from Thailand’, PIER Discussion Paper No. 20, February.

Mumtaz, H., and P. Surico (2015), ‘The transmission mechanism in good and bad time’, *International Economic Review*, 56(4): 1237-1260.

Nelson, J. C. Perfetti, D. Liben, and M. Liben (2011), *Measures of Text Difficulty: Testing their Predictive Value for Grade Levels and Student Performance*. Student Achievement Partners.

Praet, Peter (2014). ‘Current issues and challenges for central bank communication’, Speech at the conference on ‘The ECB and Its Watchers XV’. 12 March 2014.

Reis, R. (2013), ‘Central bank design’, *Journal of Economic Perspectives*, 27(4): 17-44.

Sheehan, K. M., I. Kostin, Y. Futagi, and M. Flor (2010), ‘Generating automated text complexity classifications that are aligned with targeted text complexity standards’, ETS Research Report No. RR-10-28. Princeton, NJ: Educational Testing Service.

Sheehan, K. M., I. Kostin, Y. Futagi, and M. Flor (2014), ‘The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment’, *The Elementary School Journal*, 115(2): 184-209.

Sheehan, K. M. (2015), ‘Aligning TextEvaluator scores with the accelerated text complexity guidelines specified in the Common Core State Standards’, ETS Research Report No. RR-15-21. Princeton, NJ: Educational Testing Service.

Sheehan, K. M. (2016), ‘A review of evidence presented in support of three key claims in the validity argument for the TextEvaluator text analysis tool’, ETS Research Report No. RR-16-12. Princeton, NJ: Educational Testing Service.

Sherman, L. A. (1893), *Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry*. Boston: Ginn and Co.

Stein, J. C. (2014). “Challenges for Monetary Policy Communication”. Speech given to Money Marketeers of New York University, May 6, 2014.

Tetlock, P. C. (2007), ‘Giving content to investor sentiment: The role of media in the stock market’, *Journal of Finance*, 62(3): 1139-1168.

Vogel, Mabel, and Carleton Washburne (1928), ‘An objective method of determining grade placement of Children’s Reading Material’, *Elementary School Journal*, 28: 373-381.

Appendix

Central Bank (CB)	Country / Region	Periods	# Meetings
Major CBs			
FED	USA	2000-2015	128
ECB	EU	2000-2015	181
BOJ	Japan	2000-2015	228
Inflation Targeters			
Central Bank of the Republic of Armenia	Armenia	2006-2015	82
Reserve Bank of Australia	Australia	2000-2015	109
Bank of Canada	Canada	2000-2015	125
Banco Central de Chile	Chile	2000-2015	189
Banco Central de Colombia	Colombia	2007-2015	103
Czech National Bank	Czech Republic	2009-2015	55
Hungarian National Bank	Hungary	2002-2015	253
Central Bank of Iceland	Iceland	2009-2015	51
Bank Indonesia	Indonesia	2006-2015	118
Bank of Korea	South Korea	2000-2015	192
Reserve Bank of New Zealand	New Zealand	2000-2015	64
Norges Bank	Norway	2000-2015	113
Central Reserve Bank of Peru	Peru	2001-2015	182
Bangko Sentral ng Philipinas	The Philippines	2002-2015	144
National Bank of Poland	Poland	2001-2015	172
National Bank of Romania	Romania	2003-2015	99
National Bank of Serbia	Serbia	2005-2015	174
Bank of Thailand	Thailand	2000-2015	126
Central Bank of the Republic of Turkey*	Turkey	2006-2015	123

*Data for quarterly policy rates are not available from