Journal of Global Business Review: Volume 23 No. 1 January-June 2021 pp. 33-53

THE SENTIMENT ANALYSIS OF MPC MINUTES OF THE MEETINGS AND RELATIONSHIP WITH BOND YIELD AND SET INDEX

Thatree Homsirikamol[1*], Nopphon Tangjitprom[1], Thananporn Sethjinda[1]
[1]Graduate School of Business, Assumption University of Thailand,
592/3, Ramkhamhaeng 24 Rd., Hua Mak, Bangkok Thailand, 10240

Recevied: August 12, 2020
Revised: February 8, 2021
Accepted: February 08, 2021

ABSTRACT

With the increasing application of textual analysis tool to quantify qualitative data, this study aims to contribute to the academic discussion on textual analysis of the meeting minutes of the Monetary Policy Committee of the Bank of Thailand by examining the effectiveness of sentiment analysis tool in determining the net tone of the document and how such indicator relates to the common financial market indicators such as bond yields and the SET Index. By applying Valence Aware Dictionary for sEntiment Reasoning, this study found that the net tone derived from the compound values of each MPC minutes of the meeting was consistent with key economic events for both expansion (positive sentiment) and slowdown (negative sentiment). To validate relationship with Bond Yields and SET Index, Vector Auto Regressive (VAR) has been applied to validate the relationship between the net tone and 1 year (1Y), 2 years (2Y), 5 years (5Y) and 10 years (10Y) government bond yields and the SET Index. Based on VAR, the net tones were significantly correlated with 10 years bond yields. Relationship with shorter maturities and the SET Index was not supported.

Keywords: Sentiment analysis, Textual analysis, Central bank communication, MPC Minutes of the meeting

Introduction

Over the past few decades, financial research has seen significant development, particularly in the application of computational based approaches on several fields in finance and economics. The influence of financial engineering and advanced computational techniques has been well-noted for quantitative analysis such as asset pricing, risk modeling, or trading strategies to name a few (Chen, 2002). Apart from the advanced quantitative computational approaches, the past decade has seen an increasing extension towards using text-based data in some research areas. As a result, a growing number of studies have been adopting a combination of text-mining and statistical or econometric techniques to examine information or signals from texts to complement the mainstream analysis.

Historically, text-mining has been predominantly used in other disciplines such as social science and marketing, the application in finance and

[*]corresponding author: e-mail: thatree.h@gmail.com

economics research sprouted around the mid-2000. One of the possible explanations could be attributable to the evolution of communication technology, giving rise to a wealth of new text data available through internet search records and social media (Bholat, Hansen, Santos, & Schonhardt-Bailey, 2015). With such development, researchers have been keen to draw opinions and emotions expressed by individuals to enhance the analysis of certain research areas where qualitative data could offer signals or insights from the crowd.

As a result, there has been a growing number of studies, employing text-mining and statistical analysis, such as tracking market sentiments based on news, financial market data (Pang & Lee, 2008), trading (Peterson, 2016), social media feeds (Drus & Khalid, 2019) or assessment of financial market reaction to central banks' communication and announcement (Blinder, Ehrmann, Fratzscher, Haan & Jansen, 2008).

In the areas of central bank communication, scholars recognized the application of text mining as an extension to quantitative models or event study to assess communication effectiveness and market reaction (Blinder et al., 2008; Hughes & Kesting, 2014).

Several researchers have used text processing software to supplement the traditional approach of manually quantifying and coding the central bank's view (Stekler & Symington, 2016). Furthermore, there have been growing number of studies on establishing understanding of speech acts, signaling direction or intentions and linkages between such linguistic findings on subsequent action, policy implications and reactions in the financial markets (Blinder et al., 2008; Hughes & Kesting, 2014).

On the whole, empirical evidence suggests that the increasing level of central bank communication and transparency have facilitated the predictability of monetary policy decisions by reducing noises (Blinder et al., 2008, Ehrmann & Fratzscher, 2007; Hughes & Kestin, 2014). Furthermore, communications by central banks, such as official statements, reports and minute, seem to have an impact on the levels and volatilities of various interest rates and other financial market indicators. In particular, a number of studies on FOMC statements suggest that treasury yield changes were dependent on themes expressed in the minutes as well as the level of uncertainty expressed (Boukus & Rosenberg, 2006; Mazis & Tsekrekos, 2017).

Similar results have been noted among other central banks such as Bank of England (Reeves & Sawicki, 2007), ECB Australia, New Zealand and Canada (Connolly & Kohler, 2004), noting that the quantitative magnitude do vary across central banks. For instance, the effect of interviews and speeches of Bank of England was more pronounced in short-term interest rates, compared to that of ECB. For the latter, there was an impact from short-term to medium term interest rates (Ehrmann & Fratzscher, 2007).

In recent years, scholars, in Thailand, have also been applying computational textual analysis approaches to gain insights into the information and significance of central bank communication documents. In particular, Luangaram and Sethapramote (2016) studied the effectiveness of central bank communication as the predictability of short-term interest rate, monetary transmission mechanism through word fish and cross-validation with human-rating. Furthermore, Luangaram and Wongwachara (2017) applied LDA on communication documents of major central banks to determine topics, themes, tones and relationship with policy decisions (Luangaram & Wongwachara, 2017). Finally, Chatchawan (2019) applied semantic similarities and impact on volatilities of bond yields, SET index and foreign exchange rates.

To contribute to the wider discussion on the computational textual analysis in Thailand, this study humbly proposes to explore the content analysis of Monetary Policy Statements, using a lexicon and rule-based content analysis tool. The advantage of using sentiment analysis tool is that there are readily developed tools based on well-established dictionary and available as open-sources to developers' communities. Such lexicon and rule-based approach allows efficient processing of textual information with limited influence from the researchers' biases. Apart from that, the tools calculate the compound values which takes into account the combination of positive, negative and neutral words at the sentence level or word by word.

Accordingly, this study applies Valence Awareness Dictionary and sEntiment Reasoner (VADER) to extract sentiments from the communication from the Monetary Policy Committees. The textual analysis will be supplemented by LDA analysis to determine the consistency between text themes and the text intensity determined by VADER sentiment analysis. Secondly, this paper attempts to validate whether there exists any relationship between the net tone, derived from the compound values on key financial market indicators such as bond yields and the SET Index over the study period. In all, this study addresses the research objectives below.

This study aims to determine whether the sentiment value can be determined from the monthly minutes of Monetary Policy Committee (MPC) using lexicon and rule-based analysis. Moreover, this study aims to examine whether there is any relationships between the financial market variables including the government bond yields of 1, 2, 5 and 10 years maturities as well as SET Index and the net tone calculated from the sentiment analysis.

Literature review

Market reactions to central bank policy changes have been an important research area. There have been a wide range of studies examining the impact of target rate adjustments on other interest rates, as well as the stock markets, (Boukus & Rosenberg, 2006) although analysis of central bank communication has only started in the past two decades (Blinder et al., 2008). Several researchers have pointed to the shift in communication strategy by major central banks towards communication transparency and managing the market expectation (Lucca & Moench, 2015; Boukus & Rosenberg, 2006; Blinder et al., 2008).

One interesting development in literature centres on the application of various approaches of linguistic analysis to understand clarity, classify themes and signalling direction or intentions. Some commonly used approaches include Latent Semantic Analysis (Boukus & Rosenberg, 2006), Latent Dirichet Allocation (Luangaram & Wongwachara, 2017), Alceste (Stekler & Symington, 2016) and Supervised Machine Learning. As research continues to advance, increasing number of text analysis approaches have been applied to explore different characteristics of central bank communication and linkages with text expressed via other media forms.

For instance, Luangaram and Sethapramothe (2016) applied Wordfish algorithm and econometric modelling to assess the effectiveness of monetary communication, in terms of predictability of short-run policy interest rate, monetary transmission mechanism and the anchoring inflation expectation. Results reveal that the MPC statement exerts influence over the yields with longer maturities. The increase in policy rate could anchor expected inflation only in the short run.

To assess communication effectiveness, Luangaram and Wongwachara (2017) used several computational linguistics tools to measure the document readability, topics and tones. In particular, they employed advanced Flesch-Kincaid grade (FK) and ETS Text Evaluator to measure the readability and document complexity, while using Latent Dirichet Allocation to determine the topics, while using dictionary method to determine the tone. This approach has covered major central banks under the inflation targeting scheme. Their research discovered that the readability of central bank communication was for the level of advanced readers. The tone of the document relating to economic outline significantly related to the changes in policy rates decision, supporting the monetary transmission mechanism.

Apart from the preceding approaches, text sentiment analysis has been an area with rapid development and application in various disciplines. It has also been used as a tool to measure financial market sentiment. For example, Shapiro, Sudhof, & Wilson (2019) applied text sentiment analysis on FOMC to measure the central bank's objective function. On a broader scope, Aggrawal (2019) noted several sentiment measures such as Pessimism factors, Anxiety Index, or FEARS Index, which have been widely discussed and followed today.

As scholars continue to address pending research questions, future research in central bank communication is likely to incorporate advanced computational linguistic tools to address such challenges.

Bond and capital market

The reason bond yield and capital market were picked as part of the research variables was due to the staggering empirical studies between monetary policy, bond, capital markets trying to understand the impact of the policy decision. Chun, (2010), did an in-depth empirical study by comparing different models to understand expectation on bond yields as the effect of the monetary policy. Through employing term-structure the author found interest rates react with the forward-looking in-turn affect the bond-yield movement in long-term maturity. Andrian & Shin, (2008), discovered short-term interest rates direct the cost of leverage and has influence in financial intermediary balance sheets. Concluded with monetary policy and financial stability policies are intricately linked. Beckworth, Moon, and Toles (2009) proposed to using vector autoregression to understand the monetary policy shocks. They were able to discover through chronic neutrality the monetary policy was the origin of variations of corporate bond yield spreads. Their study is based on monthly frequency from 1959 until 2008, using industrial production output, 3-month T-bill and S&P composite index as part of their equation. While majority of the studies still focuses on using economic indices to be part of the main drive of the market movement. This research, however, would like to approach from the standpoint of emotion impact on the market.

Table 1 Summary of selected research on central bank communication (from author's compilation)

Authors & years	Approaches	Data	Results
Boukus and Rosenberg (2006)	LSA & correlation with macroeconomic variables	FOMC minutes 1987-2005	Themes related to economic and monetary policy were associated with changes in treasury yields during release times.
Lucca and Trebbi (2009)	Google semantic orientation & FACTICA semantic orientation score univariate & VAR	FOMC statements 1999-2008	Short-term treasuries respond to unexpected policy rate decisions, longer-dated yields react to changes in content. FOMC statements
Hansen and McMahon (2016)	LDA & FAVAR	FOMC statements 1998-2004	Central bank guidance on future interest rate are more important than economic outlook. Neither exhibited strong effect on economic variables.
Luangaram and Sethapramote (2016)	Word fish algorithm hand-coding ordered PROBIT predictive regression SVAR	MPC statements & quarterly monetary policy report 2000-2011	Influence of MPC statements over the yields with longer maturities
Luangaram and Wongwachara (2017)	Flesch-kincaid grade (FK) and ETS text evaluator LDA & dictionary method	Policy statements from 22 central banks 2000-2015	Content of central bank communication remain complex. Net tone relates to changes in policy decision. Communication complements policy transmission mechanism.
Chatchawan (2019)	Cosine similarity tone analysis with dictionary method EGARCH	MPC press releases 2010-2018	Semantic similarity reduced the volatility of 1-month, 3-month, 10-year & 15 year government securities. No impact on SET & foreign exchange.

Sentiment analysis

Sentiment analysis is a sub-set of text mining approaches, used to determine and categories opinions, feelings or emotions (Gupta & Bhathal, 2018), expressed in various sources of texts such as social media or news. With the massive growth of information over the past few decades, sentiment analysis has become an increasingly popular tool, with applications across various domain from politics, scientific application, healthcare, telecommunication, risk management to various fields in business management such as marketing, management (Gupta & Bhathal, 2018; Qazi, Raj, Hardaker & Standing, 2016) and finance and investment (Ribeiro, Araujo, Gonclaves, Gonclaves & Benevenuto, 2016). Corresponding to the growing need for sentiment analysis, there have been continued growth in sentiment analysis methods from different fields of computer science. Broadly speaking, sentiment analysis could be grouped into two main categories: the machine learning and the lexicon-based methods. With machine learning approaches, the data is processed under supervised classification approaches with labelled data to train classifiers. On the other hand, the lexical method makes use of the existing lists of words, created to support text domains the tools are originally developed for (Ribeiro et al., 2016).

As to which methods are superior to others, there has not been a clear answer to this question. In general, the available sentiment analysis techniques are acceptable by the research community. It is also common to see applications of sentiment analysis on the same topic, but based on completely different methods, whether it be LIWC, SENTINET or VADER to name a few (Ribeiro et al., 2016). In the context of finance and economics, there have been a few examples of research using lexical approaches. For instance, Loughran & McDonald (2011) constructed their own dictionary, using data from 10 K report with five additional word lists. This dictionary has been used in several works such as Shapiro, Sudhof and Wilson (2019) on the meeting transcripts of FOMC. Other examples include Harvard GI lexicon, which is often used in conjunction with that of Loughran and McDonald (2011) for measuring negativity in news articles about companies and impact on their stock prices (Heston & Sinha, 2016).

Recent lexicon methods of sentiment analysis are developed to better capture the contextual characteristics of words within the corpus of interest. One such example is Valence Aware Dictionary for sEntiment Reasoning (VADER). According to Shapiro, Sudhof and Wilson (2017), VADER is a sentence level sentiment classifier, consisting of a list of several thousand words labelled in according to the most negative to the most positive. Furthermore, VADER takes into account a set of heuristic rules that account for a word's context within the sentence in providing the final compound values.

The subsequent section briefly discusses the development steps of VADER and recent performance assessment in comparison to other sentiment analysis approaches.

Valence aware dictionary sentiment reasoning

Developed by Hutto and Gilbert (2014) Valence Aware Dictionary for sEntiment Reasoning has been developed through a combination of qualitative and quantitative methods to produce sentiment lexicon that is suitable for analysis of social media texts and applicable to generalize across other domains. The initial lexicon to develop VADER has based on screening the sentiment word backs from well-established lexicon such as Linguistic Inquiry and Word Count (LIWC), Harvard General Inquiry dictionary (GI), and crossed validated with ANEW, SentiwordNet and SenticNet in developing valence score for sentiment intensity.

To account for valence score for sentiment intensity, the scale of intensity for positive and negative valence was based on -4 to 4. To validate the developed lexicon, machine learning and validation performed by human raters were performed on various types of texts before establishing the ground-truth of sentiment intensity corpus. As part of the validation step, testing of VADER was performed against six other well-established sentiment lexicon, on four databases covering social media text, product review, movie review and news editorial. Statistical results from the validation showed that VADER performed among the top three lexicons.

Apart from the authors' performance assessment, VADER had been included an extensive benchmark test performed by Ribeiro et al. (2016). With the aim to establish the benchmark for performance comparison of sentiment approaches, Ribeiro et al. (2016) assessed the performance of 24 lexicon based sentiment analysis methods, including open-sourced and paid software across 18 data sets, covering news comments, reviews and tweets. Their assessment revealed that some sentiment analysis tools might perform well in the database that were originally developed for but perform poorly in others.

With regards to performance of VADER, the study revealed that VADER achieved the best performance for predicting positive, negative, and neutral texts, although the tool did not achieve the first ranking across all databases. Having said that, VADER performance was relatively consistent, achieving performance in the top seven across all datasets.

At present, similar validation efforts have been performed by researchers. The additional validation by Sohangir, Petty and Wang, (2019), was performed on Stock Twits data. The performance experiment was performed against Machine Learning, Textblob, Sentiwordnet and VADER. Statistical results revealed that VADER was able to outperform Machine Learning and other Lexicon based methods.

Within the scope of this study, VADER is chosen as a sentiment analysis tool for MPC minute of the meetings on the following basis. Firstly, the tool is available as open-source library, allowing access and contribution from research community. Hence, it is easy to operate. Secondly, as the tool is fairly recently developed, applying VADER will allow deeper understanding into its strengths and weaknesses, enabling further contribution to improve the tool.

There is, however, one important caution on the use of VADER. Since VADER lexicon has originally been developed for social media text, its performance may be questionable for economics and financial corpus where profound understanding is required. To address this potential downside, the financial sentiment wordlist was updated to VADER lexicon for this research.

Methodology

Data collection

Data used in this study contains the monthly minutes of the monetary policy committee meeting in English format from January 2011 until June 2019, or 68 reports in total from the Bank of Thailand. Financial data, covering 1 Year, 2 Year, 5 Year and 10 Year government bond yields and the SET Index are obtained from iBond and Thomson Reuter Eikon, respectively.

Prior to performing sentiment analysis, the data from the minutes of the meeting is pre-processed in line with the pre-processing steps, commonly adopted by academic works in content analysis of central bank communications (Kijsriopak, 2016; Mazis & Tsekrekos, 2017). The pre-processing is performed with Natural Language Processing Tool Kits, an open library application programming interface (API).

1. Eliminate text formatting such as empty spaces, capitalization, symbols, digits or other special characters that might appear in the text.

2. Remove common and function-neutral words that do not contribute to the

information value of the document or stop words.
3. Tokenize the string of words by splitting the entire text documents into smaller units such as words or terms, which are referred to as tokens in Natural Language Processing.
4. Combining specific words for its to be recognizable in SA process, e.g. “pick ups”, “in line”.
5. Apply stemming to all the tokenized words. Stemming is the process of producing common etymological roots for words matrix.
6. Re-iterate the process for refined results (Zhao, 2018).

Figure 1 shows the details of the total words and unique words count before and after pre-processing steps.

Stopwords (SW)		**Stemming (SM)**		**Total words count**		
Words	Frequency	accelerate	acceler		Total counts	
the	15.76%	accommodative	accommod		Mean	Std.
to	9.52%	addition	addit	Before SW	1893	779
in	8.10%	advanced	advanc	After SW	1016	418
of	7.75%	appreciating	appreci	After SM	1016	418
and	7.63%	appreciation	appreci			
a	3.04%	appropriate	appropri		Unique counts	Words
as	2.34%	become	becom		Mean	Std.
that	1.89%	becoming	becom	Before SW	565	123
)	1.67%	beginning	begin	After SW	471	111
was	1.64%	capital	capit	After SM	408	85
(	1.61%	challenges	challeng			
with	1.44%	concerns	concern			
		conditions	condit			
		continually	continu			
		continued	continu			

Figure 1 Showcase stopwords, stemming and word counts (from author’s calculation)

Given that the original lexicon of VADER is developed to cater for social media tweets, this study performed an additional step to update VADER lexicon with financial sentiment words, present in the existing corpus. Researcher performed data labeling of the corpus with over 1,300 unique counts. Process sentiment analysis through VADER based on the initial results, lists of positive, negative, and neutral words were reviewed and cross-validated against the manual data labelling. Update VADER corpus based on the words that were initially identified as neutral when they should fall into positive or negative category.

Data analysis

After the pre-processing and lexicon updates were completed, the corpus was processed via VADER version no. 3.2.1. The program returned the compound values for each sentence within the minute of the meeting with the breakdown of positive, neutral, and negative scores. Sample results are shown in Table 2.

Table 2 Examples of compound values generated by VADER (from author's calculation)

Sentence	Compound	Sentiment
Result challenge Asia going forward likely come capital flow volatility appropriate pacing monetary tightening.	-0.4019	Negative
Monetary sector private credit expanded together overall economic growth.	0.6369	Positive
Commercial bank rapidly raised deposition rate following rate hike.	0	Neutral

To assess the relationship between the sentiment values and financial variables, this study calculates a time-series values of net tones for each minutes of the meeting. This study follows the calculation approach by Hansen and McMahon (2016) to compute the net tone. Net Tone equation is written as follows:

$$Net\ Tone = \frac{\sum_{i=t}^{n} SA\ Positive_t - \sum_{i=t}^{n} SA\ Negative_t}{\sum_{i=t}^{n} SA\ Sentence_t}$$

To improve the robustness of the data and nullifying the unit of measurement to compare observations measured with different units; by transforming the data into new scores with a mean of 0 and a standard deviation of 1. Z-score Normalization method was implemented to common scale the dataset without distorting the differences in the ranges of values. The NT formula can be generalized as:

$$Z - Score\ Normalization_t = \frac{(\text{Net Tone}_t - \overline{Net\ Tone})}{\sqrt{\frac{\sum_{i=1}^{n}(Net\ Tone - \overline{Net\ Tone})^2}{n-1}}}$$

To validate the compound values from VADER and the net tones, this study performs word clouds analysis for each MPC meeting minute. Word Cloud is a visualization presentation of words frequency. The most commonly terms, used in the text being analyzed, appear in word cloud based on text sizes.

Word clouds are generated based on Word cloud version 1.5.0, open library API (Zhao, 2018). Word clouds are generated from two sets of data, which are the data for each minute of the meetings and the data for all minutes during the study period. Results from word clouds will be used to support the validation of the net tones to determine the level of consistency, captured by VADER sentiment analysis.

This paper refers to the studies of Beckworth et al. (2009), they used standard Vector Auto-Regression Model to define relevant significant of impact of their chosen variables. This research main focus will be the relationship between Net Tones, bond yields and SET index, VAR will be applied for this study.

Derived from the need to forecast structural levels of economic variables, Vector Autoregression (VAR) has been primarily used for macroeconomic forecasting (Greene, 2002). It has been recognized for flexibility and adaptability for multivariate time series analysis. VAR has been be used to describe dynamic behavior of economic and financial time series.

To assess the relationship between net tone, bond yield and the SET index, this research proposes to apply VAR to determine the relationship. More specifically, VAR imposes fewer and weaker restriction on the equation comparing to its counterpart, structural modelling. Considering the low dimension, non-economic variables of the research between emotions and market movements, this factor created the suitability use for VAR. Another distinct feature of VAR model is the need to identify relationship between endogenous and exogenous variables became null, as all variables are considered as endogenous within VAR model (Lu, 2001). It is very much possible to impose other structural exogenous variables into the equation, although, Greene (2002) stated researchers have found simple and small-scale VARs have proven to be reliable. Gottschalk (2001) study also found there is a draw back on SVAR model with low-dimension model assumption. Since, VAR has the ability to obviate the results from contemporaneous, exogenous variables, it is very favorable as a tool to help establish linear relationship between multivariate time series (Huynh, 2019).

VAR model does have explicit rules such as time series have to be stationary and cointegrated with each other before approaching empirical study. To distinguish the relationship of the variables; if the p-value of tested variables is below 0.05, it is safe to reject the null hypothesis that variables does not have interrelationship with each other. Another method that is widely adopted by many empirical studies is the Granger Causality test to understand the if a variable Y can be Granger-caused to variable X.

The basis of VAR model is each time series variable can be explained by its own lags and other variables in an interconnected multivariate system. The mathematical formula is typically presented as

$$y_t = \mu + \emptyset_1\, y_{t-1} + \ldots\ldots + \emptyset_p y_{t-p} + \epsilon_t$$

Where, y_t is announce the date-t observations of stationary variables. p refers to the number of lags. $\emptyset$ can be denotes as matrix of autoregressive coefficient of the lag Y. ϵ_t is the error which is considered as white noise within the equation.

In this study, the two-series Vector Auto-Regression Model is used in this study. Each financial market variable is used at a time with Net Tone in the Vector Auto-Regression Model (VAR). Bond yields of 1Y, 2Y, 5Y and 10Y maturities along with SET index were picked as the financial market variables. The optimized lag for VAR model can be found within the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Final Prediction Error Criterion (FPE) along with a few other information criterion methods of tests. This research will abide to the AIC test which will be further discussed within methodology section. The Augmented Dickey Fuller (ADF) Unit Root test is tested to ensure the data stationary before further analysis. The results of all tests are presented in section. To test the data stationary purposes, ADF was chosen for the Unit Root test due to its robustness across different applications (Greene, 2002), and accommodating the serial correlation. The data is considered to be stationary if the null hypothesis can be rejected. The formula of ADF can be denoted as:

$$y_y = \mu + \beta t + \gamma y_{t-1} + \gamma_1 \Delta y_{t-1} + \cdots + \gamma_\rho \Delta y_{t-\rho} + \epsilon_t$$

Where y_{t-1} is the lag 1 of time series and Δy_{t-1} is the first difference of the series at time t-1. M and β are both constraints of 0 while ϵ_t is the error or white noise.

Granger Causality Test can be generalized when x_t have causality on variable y_t if the given previous information of *y*, along with past values *x*, enable to forecast the current value of y_t. The null hypothesis is to test coefficients of the past values in the regression equation less than 0.05 significance level. The Granger Causality Test can be normally tested in the context of linear regression models and formula can be written as:

$$X_{1,t} = \sum_{j=1}^{p} \emptyset_{11,j} X_{1,t-1} + \sum_{j=1}^{p} \emptyset_{12,j} X_{2,t-1} + u_{1,t}$$
$$X_{2,t} = \sum_{j=1}^{p} \emptyset_{21,j} X_{2,t-1} + \sum_{j=1}^{p} \emptyset_{12,j} X_{2,t-1} + u_{1,t}$$

Where p is the maximum number of lagged observations included within Model and ∅ is the coefficient of the model. u is the residuals error of each time series. The variable X_1 is said to Granger Cause X_2 if the lagged value of X_1 can significantly predict the value of X_2. The chi-square test is used to examine the significance of coefficients of all lagged of one variable on another variable.

Results

Overview of corpus

The study period spanned 2011 to early 2019, under two BOT Governors, namely Dr. Prasarn Trairatvorakul with the period in office from 1st October 2010 to 30th September 2015 and Dr. Veerathai Santiprabhob with the period in office from 1st October 2015 to present. This study examined the general characteristics of minutes of the meeting by performing unique word counts. Figure 2 showed that words count significantly increased during 2017, during which Thailand had seen significant development in Fintech and investment in Thailand tech start-up ecosystem (Techsauce team, 2018).

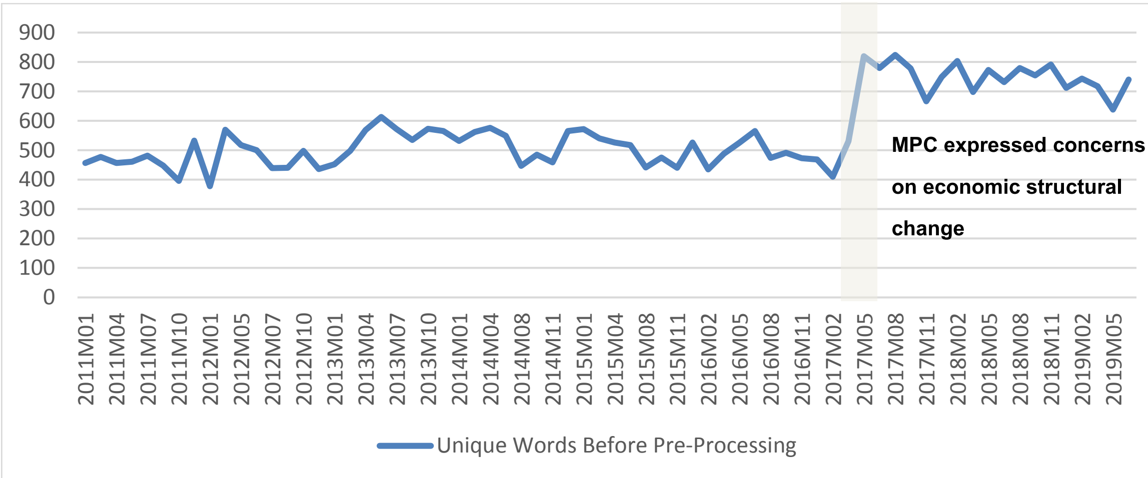


Figure 2 Unique words count (from author's computation)

During this period, the Monetary Policy Committee had expressed their view on the structural change observed in the Thai economy, and voiced their concerns on the effectiveness of monetary policy in driving economic growth as had been done historically. The additional scopes of MPC's discussion in light of this issue and development in global economies contributed to the increases in number of unique words counts post 2017.

Apart from word count, the general appearance of texts in the minutes of the meetings is obtained from word cloud analysis, Figure 3. Based on the word cloud analysis for the combined corpus, the high frequency texts reflected discussion on domestic economic conditions namely consumption, expand, momentum, inflationary pressure, momentum. Furthermore, it reflected expression of uncertainties during the study period, as evident by such words as oil, volatility, closer monitoring, partner, slowdown, external, China.

Figure 3 Word clouds for all reports (from researchers' calculation)

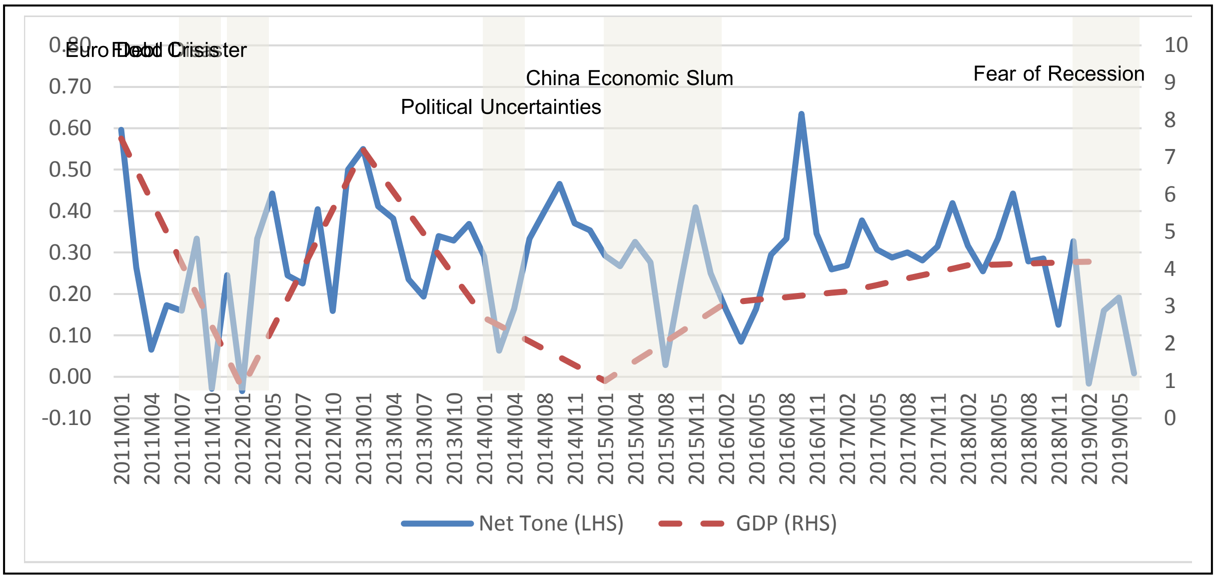
The analysis of word clouds during the study period revealed the key observations on economic activities as follows:

To assess the consistency of the net tone, the net tone values are mapped with GDP and key economic events from Table 3. Furthermore, the study reviews the themes identified through LDA analysis to identify meaning and consistency from sentiment analysis.

Table 3 Timeline of economic events (from author's compilation)

BOT Governor	Year	Economic events	Policy rates
Dr. Prasarn Trairatvorakul	2011	Strong tourism and private consumption at the beginning of the year Flood Inflationary pressures Euro zone debt crisis	2.00% p.a. 2.25% p.a. 2.50% p.a. 2.75% p.a. 3.00% p.a. 3.25% p.a. 3.50% p.a.
	2012	Euro zone debt crisis Weakened global economy Economic recovery from flood	3.25% p.a. 3.00% p.a. 2.75% p.a.
	2013	Fiscal stimulus Improvement in global economic recovery Domestic recovery led by domestic consumption and private investment	2.75% p.a. 2.50% p.a. 2,25% p.a.
	2014	Recovery in global economies, China & Asian economies Political unstability Impact on tourism, domestic demand	2.25% p.a. 2.00% p.a.
	2015	- Slowdown in Chinese and Asian economies - Financial stability - Global market volatility	1.75% p.a. 1.50% p.a.
Dr. Veerathai Santiprapob	2017	Uncertainty in global economy Volatile capital flow Gradual economic recovery	1.50% p.a.
	2018	Sovereign Debt Rising Inflation US Trade Protectionism	1.75%
	2019	Trade war between U.S. & China Drought Political uncertainty Economics slowdown	1.75%

Figure 4 shows the graphical presentation of net-tone in comparison to GDP and mapped with key economic events. The graph highlighted that the movement of net tone had been coinciding with the general economic climate during the study period. More specifically, the drop in net tone reflected slower domestic economy, caused by the great flood in 2012, as well as Thailand political instability during 2014. In view of the linkage to global economic conditions, the declines in net tone were consistent with the period of eurozone debt crisis during 2012, Chinese economic slowdown during 2015 as well as the concern on international trade conflicts between U.S. and China during 2018-2019.



Note Annual GDP’s boundary is adjusted for better comparison

Figure 4 Net tone performance

Descriptive statistics

Table 4 Descriptive statistics

	Net tone	Bond 1Y	Bond 2Y	Bond 5Y	Bond 10Y	SET Index
Observations	68	68	68	68	68	68
Mean	0.279	2.100	2.259	2.620	3.050	1439.280
STD.	0.139	0.668	0.684	0.660	0.615	218.665
Min	-0.034	1.348	1.420	1.537	1.819	938.190
Max	0.635	3.577	3.880	3.918	4.109	1792.090
Kurtosis	0.452	-1.019	-1.127	-1.481	-1.157	-0.493
Skew	-0.138	0.615	0.461	0.130	-0.010	-0.538
Durbin Watson	1.40	1.32	1.60	2.10	2.20	2.07

According to Table 4, Kurtosis for all data is within acceptable +3/-3 boundary, however there is a slight positive outlier in Net Tone while other data set are showing negative. The skewness results show the data set is closely asymmetrical made this set of data to be comfortable with its VAR model. Durbin Watson test was also carried out to test for any considerable cause of concern for autocorrelation, there appears to be small margin of positive autocorrelation for NT, 1Y and 2Y, otherwise 5Y, 10Y and SET index can be considered as having no autocorrelation.

Statistical test

Table 5 Augmented Dickey-Fuller test

	Net Tone	Bond 1Y	Bond 2Y	Bond 5Y	Bond 10Y
Test statistic	-3.9654	-5.7434	-6.5503	-8.5772	-5.0288
Critical value (5%)	-2.909	-2.909	-2.909	-2.909	-2.909
Reject H_0	Yes	Yes	Yes	Yes	Yes

The Augmented Dickey-Fuller Test or ADF test is used to examine the presence of unit root of each data series and the result has been reported in Table 5. The null hypothesis of ADF test is that the unit root is present in the data series. The critical value at 5% is -2.909. All data series have test statistic beyond the critical value at 5% significance level. This can be concluded that the null hypothesis that the unit root is present can be rejected. Therefore, all data series are stationary and ready for further analysis.

Table 6 Granger causality test

Financial market variables	Net tone granger cause financial market variables
BOND 1Y	0.192
BOND 2Y	0.137
BOND 5Y	0.005*
BOND 10Y	0.033*
SET INDEX	0.314

*indicates significant at 5% level

The result of Granger Causality Test is reported in Table 6. The number in the table is the p-value of Chi-square test that the lagged value of Net Tone can significantly predict the value of each financial market variable. In another word, if it is statistically significantly, it can be concluded that Net Tone Granger Cause that financial market variable. Based on the result, Net Tone Granger Cause only 5-year bond yield and 10-year bond yield. Therefore, this shows that the sentiment of MPC report measured by Net Tone can predict the subsequent intermediate bond yield like 5-year and 10-year. It cannot predict short-term bond yield like 1 year or 2 year as well as the stock market movement.

Therefore, the 5-year bond yield is selected to perform further analysis in Vector Auto-Regression Model (VAR). Table 7 indicates the result of significant finding of VAR model between Net Tone and 10-year bond yields. The lag length selection using Akaike Information Criterion (AIC) shows that the appropriate number of lag in VAR model is 6 lags.

Table 7 Result of VAR analysis

	Net tone	**Bond 5Y**
Constant	-0.0565	-0.0304
	(-0.483)	(-1.301)
L1. Bond 5Y	0.4873	-0.0046
	(0.663)	(-0.031)
L1. Net tone	0.2221	-0.0517
	(1.467)	(-1.711)
L2. Bond 5Y	-0.0815	0.1597
	(-0.121)	(1.184)
L2. Net tone	0.1055	0.0067
	(0.759)	(0.242)
L3. Bond 5Y	0.0573	-0.0554
	(0.091)	(-0.441)
L3. Net tone	0.0144	-0.0286
	(0.108)	(-1.071)
L4. Bond 5Y	-1.6329*	-0.0156
	(-2.659)	(-0.127)
L4. Net tone	0.0439	0.0564*
	(0.316)	(2.031)
L5. Bond 5Y	-1.4707*	-0.2662*
	(-2.219)	(-2.011)
L5. Net tone	-0.0069	0.0614*
	(-0.048)	(2.170)
L6. Bond 5Y	0.7983	-0.1663
	(1.184)	(-1.234)
L6. Net tone	0.0795	-0.0431
	(0.578)	(-1.569)

The number in parenthesis is t-stat. * indicates significant at 5% level.

From Table 7, the result of Vector Auto-regression Model shows that lag 4 and lag 5 of Net Tone can significant explain 5 year bond yield. This result is consistent with the result of Granger Causality Test reported earlier. Moreover, the Vector Auto-regression Model has been performed between net tone and other financial market variables including 1 year bond yield, 2 year bond yield, 10 year bond yield, and SET Index return. The results have not been tabulated but none of them shows any statistically significant result.

The above result is consistent with the finding of Chatchawan (2019) and Mazis and Tsekrekos (2017), where their findings stated a significant vulnerability with the long-term maturities bonds. This study finds that the relationship between Net Tone and 5 year bond yield at the 4^{th} lag and the 5^{th} lag. This can be concluded that Net Tone can represent the sentiment of MPC report and can help in predicting the subsequent intermediate-term interest rate movement.

Discussion

This study performed sentiment analysis on the minutes of the meeting of the Monetary Policy Committee during 2011-2019. Using Valence Aware Dictionary for sEntiment Reasoning (VADER), the sentence level sentiment value is used to calculate the net tone for each MPC minute of the meeting. Results are validated against word frequency through word cloud, text themes developed through Latent Dirichlet Allocation (LDA) and mapped with GDP and key economic events during the study period.

The results showed that the net tone generated from sentiment analysis was consistent in direction with economic events. In particular, the decline in net tone were consistent with slowdown in domestic economy triggered by local causes such as the great flood or the political instability, as well as the spillover effect from instability in international economies such as Euro Debt crisis, slowdown in Chinese economies or the recent trade conflicts between U.S. and China. Similar observations were also draw for the rise in net tones, which occurred in tandem with the increase in GDP and recovery of domestic economy from the implementation of fiscal and accommodative policies to support the economy.

With respect to the relationship between Net Tone with government bond yields of short-term and long-term maturities and the SET index. Results showed that the relationship between net tone and 5 year bond yields are statistically significant, based on the significant results found in Granger Causality test and Vector Auto-Regression Model (VAR). This can be concluded that Net Tone can represent the sentiment of MPC report and can help in predicting the subsequent intermediate-term interest rate movement. Likewise, there was no significant relationship between the net tone and the SET index.

The research process and results from this study offer a few key insights. Firstly, qualitative information in terms of written text does contain valuable information. This observation is consistent with the existing empirical works (Mazis & Tsekrekos, 2017; Luangaram & Sethapramote, 2016; Luangaram & Wongwachara, 2017). Secondly, there is no 'one-size-fits all' solution when it comes to textual analysis. While VADER offers intuitively easy approach to extract sentiments from the text, deeper understanding of the themes based on text occurrence in the corpus and heuristic rules are best supported by other textual analysis tools. Finally, the choice of sentiment analysis, purpose of analysis and the context of lexicon will be critical questions for researchers as more performance benchmark is being performed (Ribeiro et al., 2016). A critical implication for this observation will be increasing development of domain specific corpus as research interests continue.

Regardless of the results achieved from the research, it is important to note a few limitations of the study. First, the statistical analysis was performed solely through VAR. Consideration should be made to apply additional statistical models to verify robustness of the results. Secondly, the sentiment analysis score computed by VADER may not capture the domain specific texts, regardless of the updates to Lexicon that had been performed.

With respect to future studies, the author humbly proposes the future areas of potential research. First and foremost, important research questions such as how central bank communication influence market expectation and the public have been noted by scholars (Blinder et al., 2008). In this respect, using sentiment analysis to extract information from news might be an avenue to provide deeper understanding on expectation created by various forms of central bank communication. Secondly and most importantly, textual analysis will most likely branch to analysis performed in local language in the future. As noted by Huang, Wu and Yu (2019) in their work on textual analysis in China financial markets, the structure and expression of words in local language are significantly different from the English texts in the established lexicon. As in the case of China, textual analysis of financial text remains in its infancy.

Turning to Thailand, several research works were conducted to develop understanding on computational approaches for Thai language. For instance, Ayutthaya and Pasupa (2018) studied the implementation of Bidirectional Long short-term memory and Convolutional Neural Networks combined with part-of-speech (POS) and semantic features to analyze Thai children stories. In social media space, Vateekul and Koomsubha (2016) applied Deep Learning, namely Long Short Term Memory (LSTM) and Dynamic Convolutional Neural Network (DCNN) on Thai Twitter data with experiment results outperforming the classical machine learning based approaches.

With the continual advancement in computational language processing tools and techniques, research on textual analysis will continue to see the future application of newly developed tools for deeper level of understanding of text patterns and hidden meanings, with the combination of qualitative and quantitative tools to address research challenges.

References

Aggrawal, D. (2019). Defining and measuring market sentiments: A review of the literature. *Quantitative Research in Financial Markets,* 2-15. Retrieved from DOI: 10.1108/QRFM-03-2018-0033

Ayutthaya, T. S., & Pasupa, K. (2018). Thai sentiment analysis via bidirectional LSTM-CNN model with embedding vectors and sentic features. In *2018 International joint symposium on artificial intelligence and natural language processing (iSAI-NLP)* (pp. 1-6). n.p.

Beckworth, D. P., Moon, K., & Toles, J. H. (2009). Monetary policy and corporate bond yield spreads. *Applied Economics Letters*, *17*(12), 1139-1444.

Bholat, D., Hansen, S., Santos, P., & Schonhardt-Bailey, C. (2015). Text Mining for central banks. *Centre for Central Banking Studies,* No. 33, 1-19.

Blinder, A., Ehrmann, M., Fratzscher, M., de Haan, J., & Jansen, D. J. (2008). Central bank communication and monetary policy: A survey of theory and evidence, *Journal of Economic Literature, 46*(4), 910-945.

Boukus, E., & Rosenberg, J. V. (2006). The information content of FOMC minutes. *SSRN Electronic Journal*, 1-10. Retrieved from https://doi.org/10.2139/ssrn.922312

Chatchawan, S. (2019). *Words matter: Effects of semantic similarity of monetary policy committee's decision on discussion papers No. 121).* Retrieved from https://www.pier.or.th/wp-content/uploads/2019/12/pier_dp_121.pdf

Chen, S. H., & Kuo, T. W. (2002). Evolutionary computation in economics and finance: A bibliography. *Evolutionary computation in economics and finance Physica, Heidelberg,* 419-455.

Connolly, E., & Kohler, M. (2004). *News and interest rate expectations: A study of six central banks*, *conference-2004*. Retrieved from https://www.rba.gov.au/publications/confs/2004/connolly-kohler.html

Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *ProcediaComputerScience, 161,* 707-714.

Ehrmann, M., & Fratzscher, M. (2007). Communication by central bank committee members: Different strategies, same effectiveness? *Journal of Money, Credit, and Banking*, *39*(2/3), 509-541.

Gottschalk, J. (2001). An introduction into the SVAR methodology: Identification, interpretation and limitations of SVAR models. In *Kiel working paper No. 1072, Kiel Institute of World Economics* (pp. 1-46). Retrieved from https://www.econstor.eu/bitstream/10419/17887/1/kap1072.pdf

Greene, W. (2002). *Econometric analysis* (5th ed.). Houston: Prentice Hall.

Gupta, G., & Bhathal, G. S. (2018). *Sentiment analysis of English tweets using text mining*. Munichm, Germany: BookRix & Co.

Hansen, S., & McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, *99*(1), 114-133.

Heston, S. L., & Sinha, N. R. (2016). News versus sentiment: Predicting stock returns from news stories. In *Finance and economics discussion series.* Retrieved from http://dx.doi.org/10.17016/FEDS.2016.048

Huang, A., Wu, W., & Yu, T. (2019). Textual analysis for China's financial markets. *China Finance Review International*, *10*(1), 1-15.

Hughes, P. T., & Kesting, S. (2014). A literature review on central bank communication. *On the Horizon*, *22*(4), 328-340.

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 1-10.

Huynh, L. C. T. (2019). Spillover risks on cryptocurrency markets: A look from VAR-SVAR granger causality and student's copulas. *Journal of Risk and Financial Management*, *12*(52), 1-19.

Kijsriopak, T. (2016). *Analyzing Thai monetary policy communication1*. Unpublished bachelor's proposal, Chulalongkorn University, Bangkok.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35-65.

Lu, M. (2001). Vector autoregression (VAR)-an approach to dynamic analysis of geographic processes. *Geografiska Annaler: Series B, Human Geography*, *83*(2), 67-78.

Luangaram, P., & Sethapramote, Y. (2016). *Central bank communication and monetary policy effectiveness: Evidence from Thailand*. *PIER Discussion Paper*, 20, 1-30.

Luangaram, P., & Wongwachara, W. (2017). More than words: A textual analysis of monetary policy communication. In *PIER discussion papers No. 54, Puey Ungphakorn Institute for Economic Research.* n.p.

Lucca, D. O., & Moench, E. (2015). The pre-FOMC announcement drift. *The Journal of Finance*, *70*(1), 329-371.

Lucca, D., & Trebbi, F. (2009). Measuring central bank communication: An automated approach with application to FOMC statements. In *NBER Working Paper No. 15367, National Bureau of Economics Research*. n.p.

Mazis, P., & Tsekrekos, A. (2017). Latent semantic analysis of the FOMC statements. *Review of Accounting and Finance*, *16*(2), 179-217.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1/2), 1-35.

Peterson, R. L. (2016). *Trading on sentiment: The power of minds over markets*. New Jersey: John Wiley & Sons.

Qazi, A., Raj, R. G., Hardaker, G., Standing, C. (2016). A systematic literature review on opinion types and sentiment analysis techniques: Tasks and challenges. *Internet Research, 27*(3), 2-6.

Reeves, R., & Sawicki, M. (2007). Do financial markets react to Bank of England communication?. *European Journal of Political Economy*, *23*(1), 207-227.

Ribeiro, F. N., Araujo, M., Gonclaves, P., Gonclaves, M. A., & Benevenuto, F. (2016). SentiBench-a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science, 5*(23), 1-29.

Shapiro, A. H., Sudhof, M., & Wilson, D. (2019). Measuring news sentiment. In *Working Paper 2017-01, Federal Reserve Bank of San Francisco.* Retrieved from https://doi.org/10.24148/wp2017-01

Sohangir, S., Petty, N., & Wang, D. (2018). Financial sentiment lexicon analysis. In *2018 IEEE 12th international conference on semantic computing (ICSC)*, Laguna Hills, CA (pp. 286-289). Retrieved from doi: 10.1109/ICSC.2018.00052

Stekler, H., & Symington, H. (2016). Evaluating qualitative forecasts: The FOMC minutes, 2006-2010. *International Journal of Forecasting, 32*(2), 559-570.

Techsauce Team. (2018). *Thailand tech startup ecosystem report, 2018, investor guide, Techsauce.* Retrieved from https://techsauce.co/report/thailand-tech-startup-ecosystem-year-2018-th

Vateekul, P., & Koomsubha, T. (2016). A study of sentiment analysis using deep learning techniques on Thai Twitter data. In *2016 13th International joint conference on computer science and software engineering (JCSSE)* (pp. 1-6), Retrieved from doi: 10.1109/JCSSE.2016.7748849

Zhao, A. (2018). *NLP-in-python-tutorial*. Retrieved from https://github.com/adashofdata/nlp-in-python-tutorial