July 2023 – AN2023/3

Reserve Bank of New Zealand
Te Pūtea Matua

Analytical Notes

Words of the RBNZ: Textual analysis of Monetary Policy Statements.

Rennae Cherry and Eric Tong

Reserve Bank of New Zealand *Analytical Note Series*

ISSN 2230-5505

Reserve Bank of New Zealand
PO Box 2498
Wellington
NEW ZEALAND

www.rbnz.govt.nz

Key Findings

- Textual analysis shows that keywords mentioned in the *Monetary Policy Statements* (*MPS*) align with the objectives in the Remit.
- The tone of *MPS* has been neutral and objective, even as the sentiment in the *MPS* moves in tandem with household and business confidence surveys.
- Similar to monetary policy documents published by central banks overseas, the *MPS* are complex and may not be accessible to the general public. However, readability, which measures the complexity of a text based on sentence length and the number of syllables in words, has remained stable over 1997Q1-2021Q4 and has marginally improved recently.
- The Monetary Policy Snapshots, introduced in 2018, are easier to read than the main part of the MPS – they are accessible to a high school graduate rather than a university graduate.

1. Introduction

Clear communication helps New Zealanders understand monetary policy and its relationship to them. Communication explains to the public the purpose and rationale behind monetary policy decisions and, when done right, may enhance monetary policy transmission via different channels (RBNZ, 2020; Blinder et al. 2008; Blot and Hubert, 2018).[1] With this motivation, we apply textual analysis to flagship publications of the Reserve Bank—with the aim of assessing Reserve Bank communications and supporting its mandates of maintaining price stability over the medium term and supporting maximum sustainable employment.

The Reserve Bank reaches out to a broad spectrum of New Zealanders through different channels. These include monetary policy announcements, press conferences, business visits, testimonies to the Parliament, Governor's speeches to the public, online platforms, and laying out monetary policy actions and the reasoning behind them in the quarterly *Monetary Policy Statements (MPS)*.[2] These statements aim to 'educate and advise economic analysts and others who are providing advice to traders in the financial market on recent developments and expectations for future economic developments' (RBNZ, 2020, p.8). To determine whether the *MPS* are fit for purpose, we apply textual analysis to the *MPS*.

Textual analysis is the process of transforming text into a structured format to identify meaningful patterns and new insights (IBM, 2020). The method is increasingly used by global central banks to complement traditional sources of intelligence such as national accounts and surveys (Jansen, 2011; Haldane and McMahon, 2018; Ferrara and Angino, 2022). An advantage of textual analysis is that computer-enabled approaches can process and summarise text more efficiently than humans. Also,

[1] They include, broadly, clarifying the Reserve Bank's response under different circumstances and strengthening transmission through different asset classes. See also Guthrie and Wright (2000).

[2] These are published in February, May, August, and November after the announcement of monetary policy decisions made by the Monetary Policy Committee.

textual analysis may extract meaning from text missed by human readers, who may overlook certain patterns because they do not conform to prior beliefs and expectations (Herasymova, 2022).

We apply standard procedures developed in computer science and adapted to central banking studies to analyse chapters 1 and 2 of the *MPS* from 1997Q1–2021Q4. For each document, we tokenise the text—split raw character strings into individual elements—and remove stopwords, numbers, punctuation, and white spaces. We perform a 'human audit' to ensure the processing is performed properly, as in Shapiro and Wilson (2022). Having done so, we measure the focus, tone, and clarity of monetary policy communications conveyed by the *MPS*. We assess these issues with metrics borrowed from linguistics that have been applied to other social sciences and central banking studies. We crosscheck our process by (i) evaluating the correlation between text-based sentiment indices and conventional survey data; and (ii) interviewing authors of *MPS* to assess the consistency between information revealed by topic analysis and the authors' intention.

We focus on chapters 1 and 2 for four practical reasons. First, the opening chapters tend to contain the core messages and important news of monetary policy announcements. Second, financial analysts tend to digest the core messages within minutes of the *MPS* release and, as such, the material in the first few chapters is the most likely to influence market pricing. Third, while the first two chapters' titles have changed over time, their structure and content have remained stable, facilitating the application of textual analysis (whereas later chapters have evolved significantly over time). Fourth, limiting the scope enables a more robust human-auditing process to enhance accuracy further.

Parsing 100 *MPS* and about 141,000 words, we uncover three patterns regarding the focus, tone, and clarity of the Reserve Bank communication.

Regarding focus, we find that inflation-related words or phrases are mentioned most frequently, followed by employment-related words, finance-related words, and words that pertain to 'transitory effects'. These findings are consistent with the objectives of the Reserve Bank set out in the Remit. A corollary of our analysis is that the frequency of inflation-related words approximates the business cycle, suggesting that the Reserve Bank's attention to price and other objectives is time-varying and state-dependent, consistent with the flexible inflation targeting framework (RBNZ, 2020). Also, employment-related topics have gained increasing attention even before the adoption of the dual mandate in 2018. This result is consistent with the finding of Jacob and Wadsworth (2018) that the RBNZ's inflation targeting strategy was comparably flexible both before and after adopting the dual mandate.

With respect to sentiment, we derive two sentiment indices by counting the number of words that contain positive and negative sentiment determined by the Hu and Liu (2004) and Loughran and McDonald (2011) dictionaries (HL and LM henceforth). Both dictionaries have a predictive accuracy on economic texts comparable to the Harvard General Inquirer Dictionary, which is another authoritative dictionary (Shapiro et al. 2019). The LM dictionary is built for economic and financial professionals, whereas the HL dictionary is extracted from movie reviews that reflect the sentiment of the wider public.

We find that the sentiment ratios derived from the dictionaries match the confidence of New Zealand households and businesses revealed in surveys. Specifically, the HL sentiment ratio is positively and significantly correlated with the ANZ Roy Morgan consumer confidence index and, equivalently, the LM sentiment ratio is significantly correlated with the NZIER Business Confidence Index.

Overall, the Reserve Bank has used objective language and a neutral tone in the *MPS*. Fewer than 6 percent of words in *MPS* carry positive or negative sentiment as identified by the HL and LM dictionaries. Also, a 'polarity score' reveals that the *MPS* sentiment lies in the middle of the range of sentiment, suggesting that the Reserve Bank tone is largely neutral.[3]

With respect to clarity, we find that the *MPS* are fairly technical though consistent with international findings. Using readability metrics that measure the ratio of words to sentences and syllables to words in a document, we find that it generally takes the equivalent of a Bachelor's degree to understand *MPS* material. These results illustrate the importance of other channels through which the Reserve Bank communicates monetary policy, including *Monetary Policy Snapshots*, which provide an accessible means of communication for a broader audience.

In the remainder of the article, we discuss the methodology in Section 2. In Section 3, we present and discuss the results. Section 4 performs further robustness analyses, and Section 5 concludes.

2. Model

We apply topic, sentiment, and linguistic analysis to the *Monetary Policy Statements* (*MPS*) of the Reserve Bank from 1997Q1–2021Q4 to capture the focus, tone, and clarity of monetary policy communications. The dates selected reflect *MPS* availability at the starting point of the analysis. This section briefly describes the text data pre-processing and the text analysis methodology implemented.

2.1 Text data and processing

Our data is sourced from chapters 1 and 2 of each *MPS*. There are several reasons for focusing on these chapters. First, the opening chapters tend to contain the key messages and important news of monetary policy announcements. Second, financial analysts tend to digest the key messages within minutes of the *MPS* release, and as such, it is likely that market-moving information is contained in the first few chapters. Third, while the titles of chapters have changed over time, the structure and content have remained broadly stable, facilitating the application of textual analysis. Fourth, by limiting the scope, a more robust human-auditing process can be employed to enhance accuracy even further.

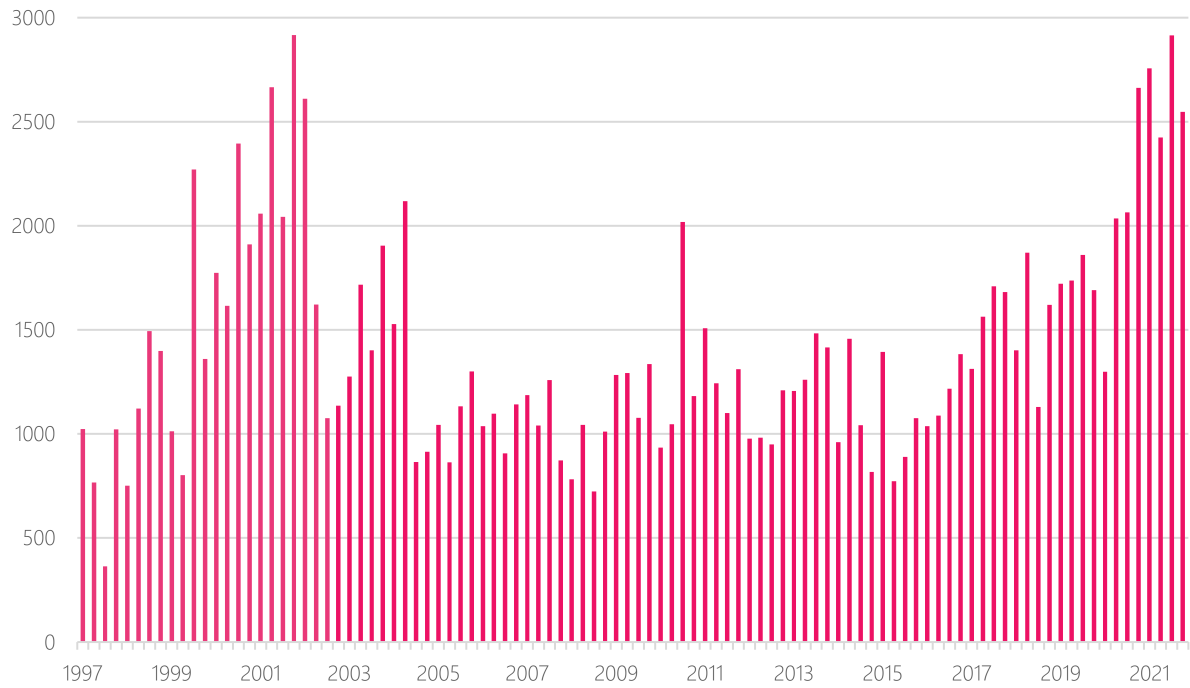
Chapter 1 covers the overall policy assessment of the monetary policy round. It summarises the messages of subsequent chapters. It is also signed off by the Governor, further sealing its official status. Since the introduction of the Monetary Policy Committee to the monetary policy framework in 2019, chapter 1 contains a summary record of meeting that records the consensus and differences within the Committee. Following chapter 1, chapter 2 describes the economic conditions of a specific monetary policy round. As the public and the central bank may at times hold different information sets—due to different assessments of the available data—chapter 2 is often the gateway for the

[3] The polarity score ranges from -1 to 1, where 1 refers to very positive and -1 very negative. *MPS* scores between -0.5 and 0.5, which is deemed as sitting in the neutral range (Hutto and Gilbert, 2014).

public to assess the information available to the central bank and how they respond to recent developments.

We clean the text in each Statement. This process includes tokenising the text, i.e. splitting raw character strings into individual elements, removing English stopwords (e.g. 'the', 'for', 'and'), numbers, punctuation, and white spaces. Text pre-processing is a common method in text analysis to reduce the data dimensionality, which is beneficial for both the computation and the interpretability of the model (Gentzkow et al. 2019). The total number of words in chapters 1 and 2 is summarised in Figure 1.

Figure 1. Number of words in chapters 1 and 2 of MPS over time



2.2 Topic analysis

We use topic analysis to investigate the focus of the Reserve Bank over time. We apply a dictionary technique, which consists of creating a list of keywords related to a specific topic and matching these words with those present in the *MPS* chapters.[4] The number of matches in each *MPS* is then divided by the total number of words of each Statement to avoid over-representing longer texts. In this way, we measure the intensity of the focus on a specific topic at the Statement level, based on the frequency of keywords for each document.

We create multiple text bags to account for different topics. To investigate related but distinct topics, we first create a list of keywords for each of the price stability and employment topics, which are the primary objectives of the Reserve Bank. To compare these primary objectives with the evolution of other topics, we create two other lists of text related to secondary objectives of the Reserve Bank, namely, finance-related topics and words that pertain to transitory effects. These words are

[4] For an application of dictionary techniques to extract the topics of central bank communication see Hansen and McMahon (2016).

consistent with Fraccaroli et al. (2020), and were verified internally by writers of *MPS* who advised on the link between word choice and the intention of the messages. The words selected for each topic are shown in Table 1. Words that share the same root but with different prefixes and suffixes are included in the analysis.

Table 1. Text bags for topic analysis

Topics	Words
Price-related	'price', 'inflate', 'cpi', 'deflate', 'tradable', 'wage', 'salary', 'pay', 'anchor', 'medium term', 'embed', 'cost', 'capacity pressure', 'supply capacity', 'supply chain', 'midpoint', 'target', 'persist', 'scarce', 'slack', 'resource'
Employment	'mse', 'participate', 'skill', 'matching', 'migrate', 'hour', 'overtime', 'employ', 'firing', 'fixed-term', 'full-time', 'part-time', 'inactive', 'job', 'labour', 'vacancy', 'worker', 'working age', 'working time'
Finance	'financial', 'banks', 'systemic risk', 'contagion', 'bubble', 'misalign', 'credit', 'insurer', 'hedge fund', 'investment fund', 'securities', 'lever', 'capital', 'derivative', 'off-balance', 'special purpose vehicle', 'payment system', 'settlement system', 'deposit', 'loan', 'currency', 'correlation', 'exchange rate', 'liquid', 'debt', 'twi', 'bond', 'swap', 'volatile', 'uncertain'
Transitory	'temporary', 'transitory', 'short-lived', 'one-off', 'look through'

Note: words with the same root and with different prefixes and suffixes are covered in the analysis. These categories are not mutually exclusive.

2.3 Sentiment analysis

We apply a similar methodology to measure the tone of *MPS*. Following the literature on sentiment analysis applied to texts, it is possible to obtain a quantitative estimate of a document's tone by matching the words with predefined lists of positive and negative terms (Loughran and McDonald, 2011; Kearney and Liu 2014).

Different from the topic analysis, in this case, we do not create our own word list, but rely on the lists of positive and negative sentiments created by Hu and Liu (2004) and Loughran and McDonald (2011). Both HL and LM are leading lexicons used by linguists to detect sentiment and are complementary to each other. They have a predictive accuracy on economic texts comparable to the Harvard General Inquirer Dictionary (GI), another authoritative dictionary (Shapiro et al. 2019). By evaluating the performance of GI, LM and HL on a database of economic and financial news, Shapiro et al. (2019) find that LM and HL lexicons assign sentiment scores similar to human ratings on the same articles, while sentiment scores of the GI lexicon are less correlated with human ratings.

Furthermore, HL and LM dictionaries are derived from different audience ratings, hence we use both lexicons as they are complementary. The LM is built specifically for the economic and financial domain relevant to the context of *MPS*.[5] In contrast, the terms in HL are extracted from a feature space of movie reviews, and hence may capture the range of sentiment of the wider public. Moreover, an additional benefit of HL is that it relies on more robust sentiment scores, as they are extracted from the rating assigned by the reviewers on their own reviews.

From the dictionaries, we compute positive and negative scores based on the count of words matched with each bag in each Statement. Once we obtained these scores, we take the difference between positive and negative terms, to get an estimate of net sentiments (Twedt and Rees, 2012). Moreover, we normalise net sentiments by the total number of words in each Statement, to prevent upward or downward bias due to longer Statements rather than the intensity of the tones. A similar sentiment ratio is proposed by Shapiro et al. (2019) and Nyman et al. (2018). Formally, for each Statement at time t, we compute the following ratio:

$$Sentiment\ ratio_t = \frac{Positive_t - Negative_t}{N_t},$$

where $Positive_t$ and $Negative_t$ are the number of terms matched in each Statement and N_t is the total number of words in each Statement. As pointed out by Shapiro et al. (2019), an advantage of this approach is simplicity and transparency.[6]

In addition, we compile the following ratios:

$$Polarity\ ratio_t = \frac{Positive_t - Negative_t}{Positive_t + Negative_t},$$

and

$$Subjectivity\ ratio = 100 \times \frac{Positive_t + Negative_t}{N_t}.$$

The polarity score is a ratio between -1 and 1. A score in the range of -1 to -0.5 typically indicates negative sentiment. A score between -0.5 and 0.5 indicates neutral sentiment, and a score in the range of 0.5 to 1 typically indicates positive sentiment. The subjectivity score captures the number of words that carries sentiment as a proportion of total number of words.

2.4 Readability analysis

We measure communication clarity using three readability metrics: the Flesch Reading Ease score (Flesch 1948), the Flesch–Kincaid Grade Level, and the Gunning-Fog Index. These metrics measure how easy it is to read a piece of text and are based on word length and sentence length. Words with many syllables are assumed to be more difficult to read than words with fewer syllables. Similarly,

[5] It uses words extracted from the annual reports that US firms submit to the Securities Exchange Commission to summarise their financial performance.

[6] In addition, they note that this approach is mathematically equivalent to assigning a score of 1 to positive matches and a score of -1 to negative matches and averaging the word-specific valence scores across all words in a text.

sentences with many words are assumed to be more difficult to follow than sentences with fewer words. Jansen (2011) uses the first two metrics to measure clarity of central bank communication, and suggests that opacity can lead to missed or misunderstood information, or even the abandoning of reading. Similarly, Ferrara and Angino (2022) use the latter two metrics to analyse the communication clarity of the European Central Bank. It is broadly agreed that the readability of central bank reports is important for the public to understand monetary policies.

The Flesch Reading Ease (FRE) Test (Flesch, 1948) indicates easiness to read the text. The index is calculated as follows:

$$FRE = 206 - 1.02\left(\frac{Total\ words}{Total\ sentences}\right) - 84.6\left(\frac{Total\ syllables}{Total\ words}\right).$$

FRE scores are higher when passages of text contain fewer polysyllabic words and shorter sentences, while they are lower when passages of text contain many polysyllabic words and longer sentences. As a general rule, the higher the FRE score, the easier it is for a reader to read the text. This is contrary to most readability scores where a lower score reflects easier readability.

The Flesch–Kincaid (FK) Grade Level derives from FRE and provides a score that can be interpreted as the number of years of education required to understand a text. The higher the score, the greater the complexity of the language used. The measure is calculated by the formula (Kincaid et al. 1975):

$$FK = 0.39\left(\frac{Total\ words}{Total\ sentences}\right) + 11.8\left(\frac{Total\ syllables}{Total\ words}\right) - 15.6$$

Unlike the FK Grade Level, the Gunning-Fog (FOG) Index considers the number of *complex words*, rather than total syllables. Complex words are defined as words with three or more syllables, and that are not proper nouns, familiar words,[7] or compound nouns. Thus, the FOG Index differs from the FK Grade Level since it relies on vocabulary-based features, namely word categories (e.g., familiar words, proper nouns, etc.) operationalised by pre-specified lists. The metric is calculated based on the following formula (Gunning, 1952):

$$FOG = 0.4\left[\left(\frac{Total\ words}{Total\ sentences}\right) + 100\left(\frac{Complex\ words}{Total\ words}\right)\right].$$

We believe these three metrics are apt to analyse *MPS* clarity. Initially developed in the field of education research, these metrics have been applied in a wide array of other professional fields, ranging from medicine (e.g., Paasche-Orlow et al. 2003) to journalism (e.g., Wasike, 2018), and including many studies in economics (e.g., Jansen, 2011; Bulir et al. 2014; Smales and Apergis, 2017) and political science (e.g., Spirling, 2016; Lin and Osnabrügge, 2018; Schoonvelde et al. 2019; Rauh et al. 2020).[8]

---

[7] Familiar words are those which readers in general easily recognise and understand because they use them on a regular basis. Familiar words are absent of jargon.

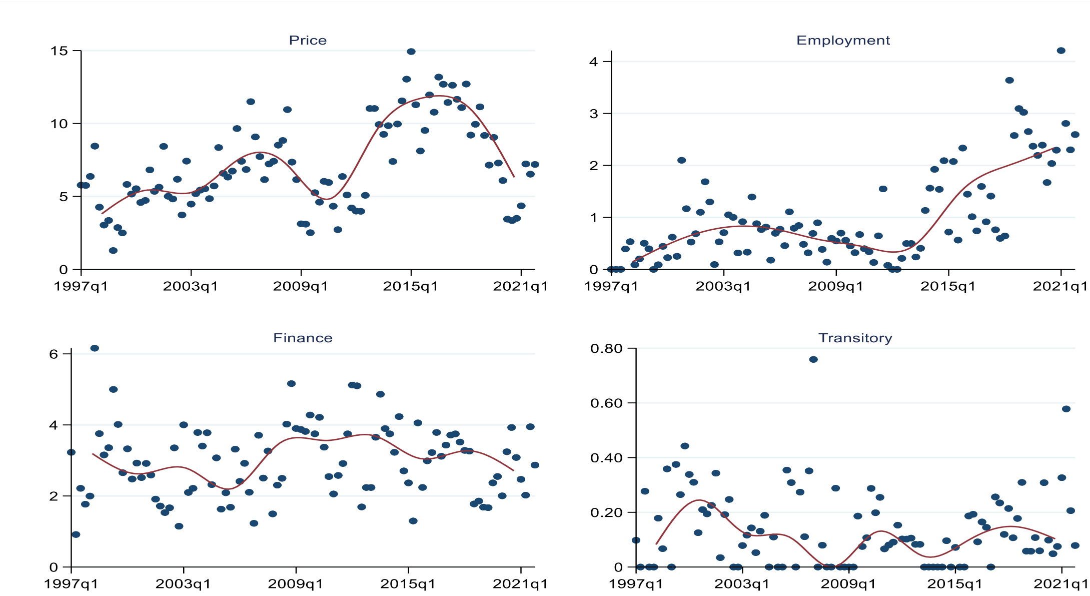
[8] The FRE Test, the F–K Grade Level and the FOG Index remain widely used in the social science literature for good reasons. For example, as Schoonvelde et al. (2019) point out, Benoit et al. (2019) find the FRE Test to be a crucial predictor of sophistication in political texts: with that measure alone, it is possible to correctly predict 72% of the human coders' judgements of difficult text snippets in their study. The introduction of various additional text features only marginally improves the prediction capacity of the FRE Test alone.

3. Results

3.1 Topic analysis

Figure 2 shows that price-related topics have been discussed most frequently, followed by finance-related, employment-related, and transitory-related topics. Notably, the attention to finance-related topics has increased since the global financial crisis (GFC) and has remained steady since, and employment-related topics have gained attention even before the introduction of the dual mandate in 2018. Transitory-related topics are least-mentioned, but this could be attributable to fewer words in the text bags and the inherent challenge in capturing the concept of being transitory. Collectively, our results show that the Reserve Bank has given regard to topics required by the Remit.

Figure 2. Percentage of topic related words in chapters 1 and 2 of MPS



Note: For each topic the graphs are scaled differently. Also, the line is a cubic spline based on cross-median values of the topic ratio scores, which are depicted by the scatter plot.

3.2 Sentiment analysis

We verify the validity of sentiment ratios using correlations between sentiment indices derived from HL and LM dictionaries and two leading indices of sentiment obtained from surveys. The ANZ New Zealand Roy Morgan (ANZ) and NZIER New Zealand Expected General Business Situation (NZIER) surveys are independent measures of consumer and business confidence and serve as effective instruments for soliciting attitudes and opinions about the current and future strength of the economy (ANZ, 2022; NZIER, 2022). Our prior expectations are that the LM sentiment will be correlated with the NZIER sentiment as both capture professional sentiment, while the HL sentiment index would be consistent with ANZ as both capture public sentiment. In fact, Table 2 reveals not only that ANZ and NZIER are correlated, but also that HL is significantly correlated with ANZ, and LM with both ANZ and NZIER. These results show that the sentiment conveyed in the text of *MPS* is consistent with the sentiment among the public at the time.

Table 2. Sentiment correlation matrix

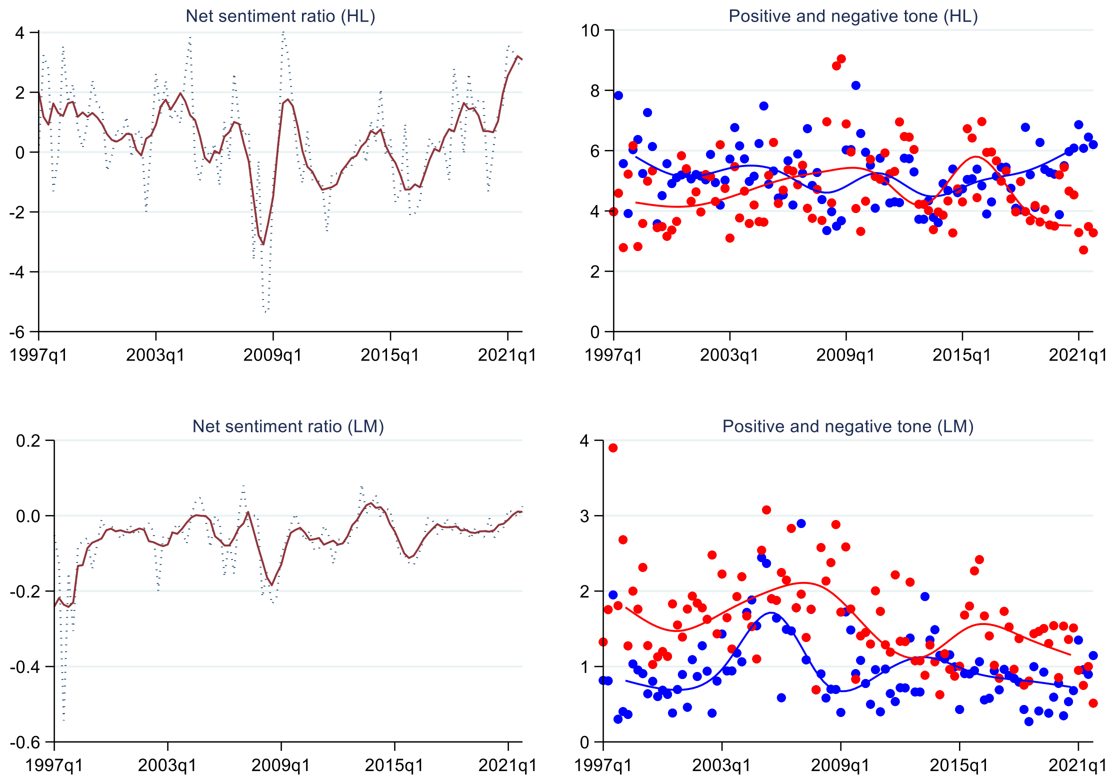
	NZIER	ANZ	Sentiment index (Hu and Liu dictionary)	Sentiment index (Loughran and McDonald dictionary)
NZIER	1.00			
ANZ	0.37**	1.00		
Sentiment index (Hu and Liu dictionary)	0.10	0.31**	1.00	
Sentiment index (Loughran and McDonald dictionary)	0.22*	0.52**	0.34**	1.00

Note: NZIER is the New Zealand Expected General Business Situation for Next 6 Months: Economy-Wide (%). ANZ is the New Zealand: Roy Morgan Survey: Overall Index (SA, Index).

The left panel of Figure 3 plots the net sentiment ratios classified by the HL and LM dictionaries. We find that sentiment is time-varying, reaching a trough around the GFC while recording higher readings in recent years. Sentiment also tends to be more volatile during crises. Without further empirical estimation, it is uncertain whether the changes in tone drive economic developments or the other way round.

While the left panel of Figure 3 displays the evolution of sentiment, it does not show whether that evolution is determined by shifts in positive or negative sentiment. Due to the structure of the sentiment ratio, lower net sentiment might be driven by an increase in negative sentiment, a decrease in positive sentiment, or both. The right panel of Figure 3 shows the evolution of two measures of positive and negative tones separately. While positive and negative tones are similarly stable on the HL measure, the LM measure instead captures more variation in negative tone. Thus, on the LM measure, negative tone has had a larger impact on net sentiment than positive tone.

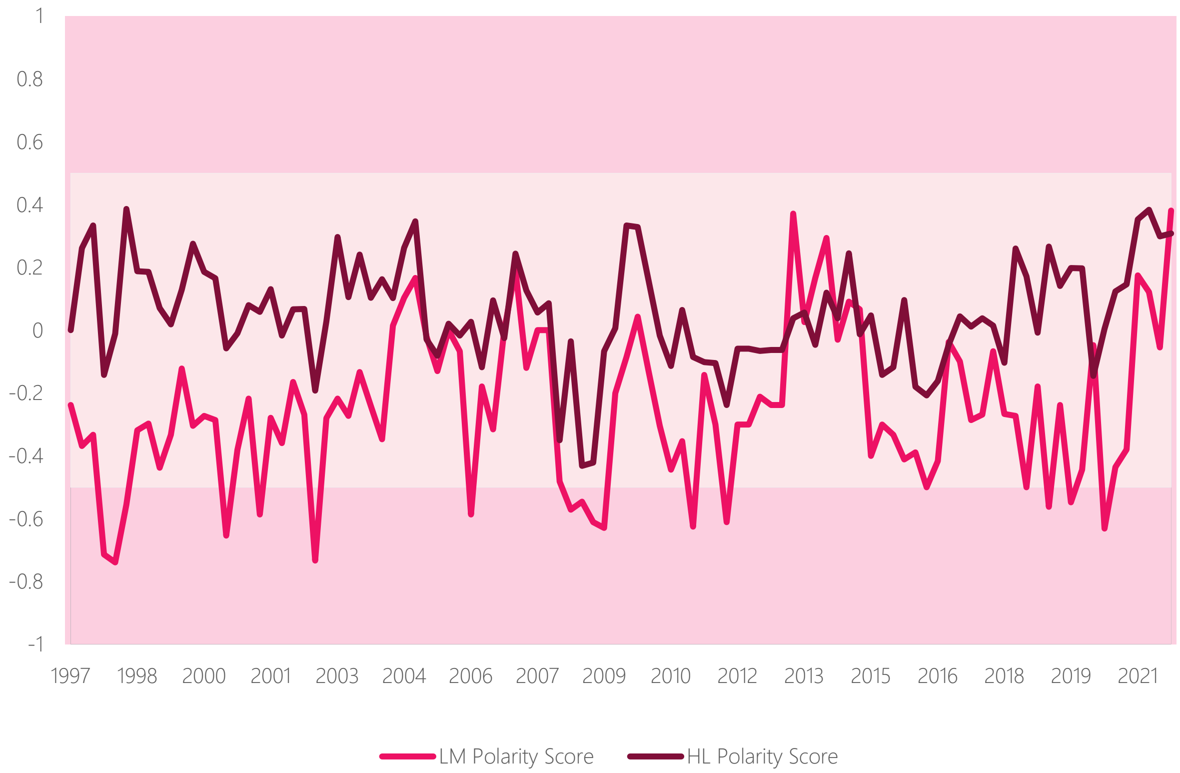
Figure 3. Sentiment tests



Note: the dashed line on the left panel is the original net sentiment ratio, and the solid line is a five-quarter moving average of the net sentiment centred at the midpoint. Gross positive sentiment is denoted as blue dots on the right panel, and gross negative sentiment is denoted as red dots on the right panel.

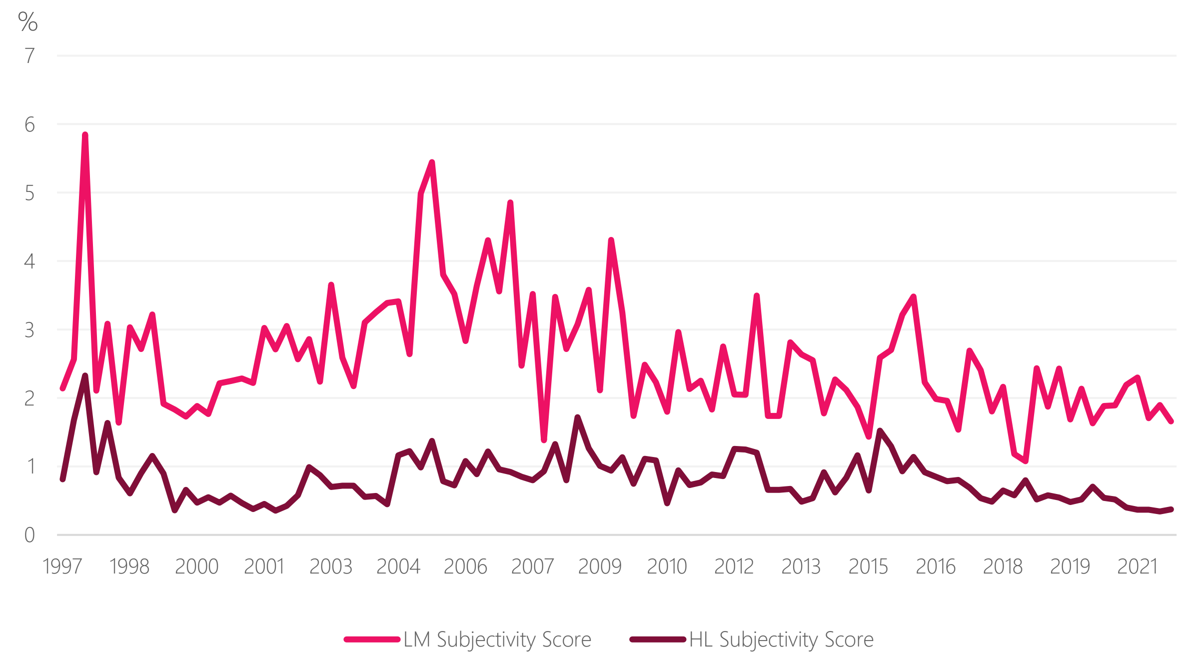
By use of the polarity and subjectivity indices, Figures 4 and 5 show that *MPS* tends to be neutral in tone and objective. Figure 4 shows that both measures of sentiment mostly lie within the range of -0.5 to 0.5 (the light pink region), indicating that the text's tone is neutral and largely devoid of positive and negative statements. Only during the Asian Financial Crisis and the Global Financial Crisis did the sentiment of the text dip into the negative range for relatively protracted periods. Similarly, Figure 5 shows that only a small proportion of words carry sentiment – fewer than 6 percent when measured by the LM dictionary and 2.5 percent by the HL dictionary. The rest of the words in the *MPS* (after cleaning) are objective in nature.

Figure 4. Polarity



Note: a score in the range of -1 to -0.5 typically indicates negative sentiment. A score between -0.5 and 0.5 indicates neutral sentiment, and a score in the range of 0.5 to 1 typically indicates positive sentiment.

Figure 5. Subjectivity

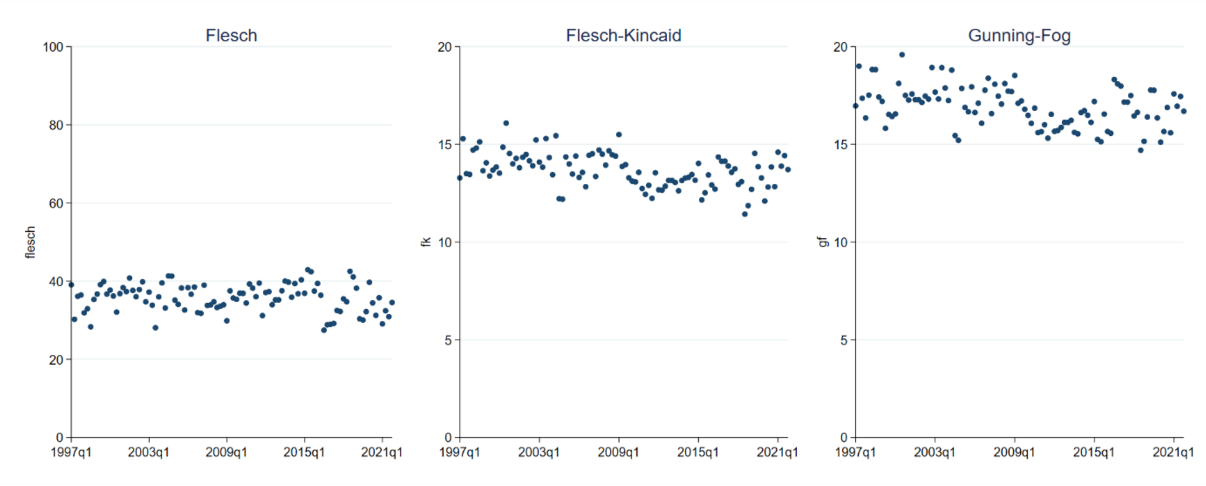


Note: the subjectivity score captures the number of words that carry sentiment as a proportion of the total number of words.

3.3 Readability analysis

To recall, the higher the FRE index, the clearer the message, and vice versa for the FK and FOG indices. Figure 6 shows that overall, *MPS* is technical and not easily digestible by the general public. The FRE scores for the MPS range from 30 to 50, which according to the original scale of Flesch (1948), means the text is 'difficult to read'. The Flesch-Kincaid grade level suggests the text is best understood by graduates with a secondary school education or above. Similarly, the Gunning-Fog index shows that it requires about 18 years of formal education—consistent with the education of a university graduate—to understand the messages in the *MPS*.These findings are consistent with the *MPS* intended audience of economic analysts and financial advisors who typically hold a tertiary qualification. These results also highlight the importance of the Reserve Bank's other monetary policy communication channels, such as speeches and social media, in reaching a wider range of audiences.

Figure 6. Readability tests

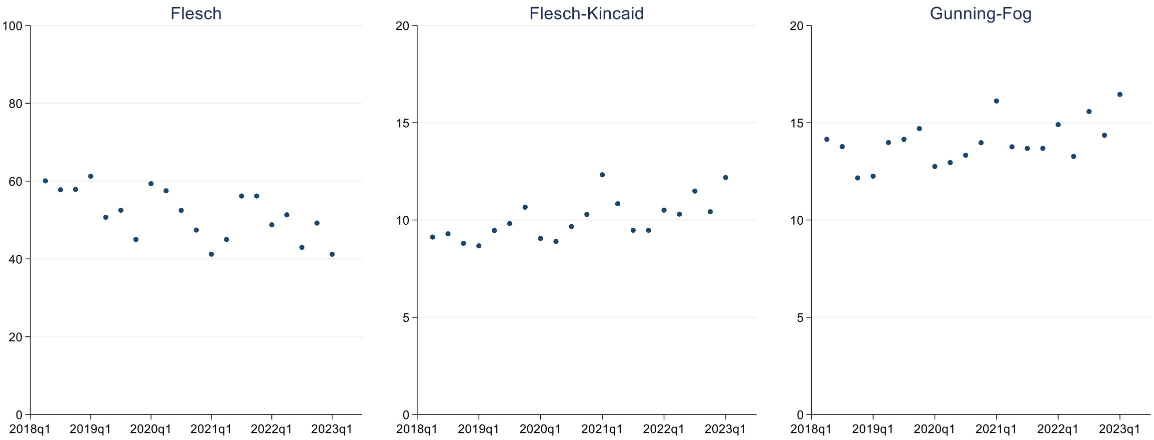


Note: The Flesch Reading Ease score (left panel) ranges from 0–100. The lower the score, the harder it is to understand a text. The Flesch-Kincaid grade level (middle panel) and the Gunning-Fog index (right panel) refer to the number of years of full-time education needed to understand the *MPS*. The higher these scores are, the harder it is to understand the text. See Section 2.4 for a detailed explanation of their interpretation.

Since the second quarter of 2018, the Reserve Bank has introduced the *Monetary Policy Snapshots* alongside the *MPS* as a supplementary resource. These *Snapshots* use straightforward language and sentence structure along with illustrations to provide a more accessible means of communication for a broader audience.

In Figure 7, we apply the same readability analysis to the text of the *Monetary Policy Snapshots*, and find a notable improvement in clarity. The Flesch Reading Ease score for the same period has risen from 40 to 50 on average, while the Flesch-Kincaid grade level and Gunning-Fog index, which gauge the years of formal education required to comprehend a text, have decreased from a range of 14–17 years (a university graduate) to approximately 10–14 years (a high school graduate). These results reflect the complementary nature of *MPS* and *Monetary Policy Snapshots*, enabling readers to choose the most appropriate material based on their individual needs and preferences.

Figure 7. Readability analysis of Monetary Policy Snapshots



Linguistic complexity is common across central bank documents. Haldane and McMahon (2018), estimate that most central bank publications have an FK grade level of 14-18, roughly equivalent to university proficiency. Based on literacy levels, Ferrara and Angino (2022) determine that *Monetary Policy Statements* of the European Central Bank, which scored an average FK grade of 14.4, were inaccessible to 90% of the general public. Similarly, readability measures classify almost every testimony by Federal Reserve Chairmen Paul Volcker and Alan Greenspan as difficult to read with FK means of 14.8 and 15.3, respectively (Jansen 2011). Among the most complex language were the statements of Henrique Meirlles and Alexandre Tombini, former Presidents of the Central Bank of Brazil, who respectively scored FK means of 18.32 and 17.83 for their communications while in office (Montes and Nicolay, 2017).

4. Robustness check

To check that the results of Figure 1 are not driven by outliers and that the evolution of words is stable over time, Figure 8 reports word clouds constructed based on the frequency of appearance of words. A larger font represents more frequent mentioning. In this exercise, we allow for bigrams to capture phrases of two words, but we have set a high bar such that only words that are likely to occur together are counted as bigrams. The threshold we use for this is a Dunning likelihood collocation score greater than 60 (Manning and Schütze, 1999).

What stands out from the figures is that similar patterns of words have been used from 1997 Q1 to 2015 Q3. During this period, topics related to inflation were relatively prominent. This is consistent with the inflation targeting framework that prevailed during this time, including the various revisions to the inflation target. However, since 2015 Q4, it is discernible that topics related to employment have received relatively more attention – the word 'employment' becomes more prominent in the final word cloud. This finding is in line with the formal adoption of the dual mandate since 2018.

Figure 8. Word clouds by time period

5. Conclusion

We analyse the Reserve Bank communication in *MPS* and find that it is consistent with the objectives set out in the Remit. The Reserve Bank has given primary attention to inflation and employment topics, with increasing emphasis on employment even before the adoption of the dual mandate. Based on these results, we infer that the Reserve Bank has been targeting inflation flexibly, facilitating both inflation and macroeconomic stabilisation.

Additionally, the Reserve Bank tone has been neutral and objective even as its sentiment moves in tandem with the public. Although the inherent complexity of the subject has resulted in a fairly technical *MPS*, which is in line with other central banks' documents and the *MPS* target audience, the overall clarity of *MPS* has remained stable over time and has marginally improved recently.

The inherent complexity of the *MPS* highlights the importance of additional channels, such as the *Monetary Policy Snapshots*, used by the Reserve Bank to reach a broader audience. The complementary nature of these communication approaches is a topic that merits much closer analysis in the future.

References

Apergis, N., and Smales, L. (2017). Understanding the impact of monetary policy announcements: The importance of language and surprises. *Journal of Banking and Finance*, *80*, 33 – 50.

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

Bulíř, A., Čihák, M., and Jansen, D. (2013). What drives clarity of central bank communication about inflation? *Open Economies Review*, *24*, 125–145.

Blot, C., and Hubert, P. (2018). Central Bank Communication during Normal and Crisis Times. European Parliament Committee on Economic and Monetary Affairs.

Ferrara, F., and Angino, S. (2022). Does clarity make central banks more engaging? Lessons from ECB communications. *European Journal of Political Economy* (Forthcoming).

Fraccaroli, N., Giovannini, A., and Jamet, J. (2020). Central banks in parliaments: a text analysis of the parliamentary hearings of the Bank of England, the European Central Bank and the Federal Reserve. *European Central Bank Working Paper*, No 2442.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221–233.

Gunning, R. (1952). The technique of clear writing. McGraw-Hill, New York.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535 - 74.

Guthrie, G., and Wright, J. (2000). Open mouth operations. *Journal of Monetary Economics*, *46*(2), 489-516.

Hu, M., and Liu. B. (2004). Mining and summarizing customer reviews. Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 168–177.

Hansen, S., and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, *99*(S1), 114-133.

Haldane, A., and McMahon, M. (2018). Central bank communications and the general public. *AEA Papers and Proceedings*, *108*, 578 – 83.

Herasymova, D. (2022). Seven benefits of using AL to Perform text analysis.

IBM. (2020). Text mining.

Jansen, D. (2011). Mumbling with great incoherence: Was it really so difficult to understand Alan Greenspan? *Economics Letters*, *113*(10), 70-72.

Jacob, P., and Wadsworth, A. (2018). Estimated policy rules for different monetary regimes: Flexible inflation targeting versus a dual mandate. *Reserve Bank of New Zealand Analytical Note*, AN 2018/11.

Kearney, C., and Liu, Sha. (2014). Textual Sentiment in Finance: A Survey of Methods and Models. *International Review of Financial Analysis*, 33.

Kincaid, J., Fishburne, P., Robert, P., Rogers, R., and Chissom, B. (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. *Institute for Simulation and Training*, 56.

Loughran, T., and McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, *LXVI*(1), 35 – 65.

Lin, N., and Osnabrügge, M. (2018). Making comprehensible speeches when your constituents need it. *Research and Politics*, 1- 8.

Manning, Christopher D. and Schütze, Hinrich (1999). *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA.

Montes, G., and Nicolay, R. (2017). Does clarity of central bank communication affect credibility? Evidences considering governor-specific effect. *Applied Economics*, *49*(32*)*, 3163-3180

Nyman, R., S. Kapadia, D. Tuckett, D. Gregory, P. Ormerod, and Smith, R. (2018). News and narrative in financial systems: Exploiting big data for systemic risk assessment. *Bank of England Working paper*, 704.

Paasche-Orlow, M., Taylor, H., and Brancati, F. (2003). Readability standards for informed-consent forms as compared with actual readability. *The New England Journal of Medicine*, *348.* 721 – 726.

RBNZ. (2020). Monetary policy handbook.

Rauh, C., Bes, B., and Schoonvelde, M. (2020). Undermining, defusing or defending European integration? Assessing public communication of European executives in times of EU politicisation. *European Journal of Political Research*, *59*(2), 397 – 423.

Shapiro, A. h., M. Sudhof, and D. Wilson (2019, June). Measuring news sentiment. *Federal Reserve Bank of San Francisco Working Paper*, 2'217-01.

Shapiro, A., and Wilson, D. (2022). Taking the Fed at its word: A new approach to estimating central bank objectives using text analysis. *The Review of Economic Studies*, *89*(5), 2768 – 2805.

Spirling, A. (2016). Democratization and Linguistic Complexity: The Effect of Franchise Extension on Parliamentary Discourse, 1832–1915. *The Journal of Politics*, *78*(1).

Schoonvelde, M., Brosius, A., Schumacher, G., and Bakker, B. (2019) Liberals lecture, conservatives communicate: Analyzing complexity and ideology in 381,609 political speeches. *PLOS ONE 14*(2).

Twedt, B., and Rees, L. (2012). Reading between the lines: An empirical examination of qualitative attributes of financial analysts' reports. *Journal of Accounting and Public Policy*, *31*(1), 1 – 21.

Wasike, B. (2018). Preaching to the choir? An analysis of newspaper readability vis-a-vis public literacy. *Journalism*, *19*(11), 1570–1587.