PPOL6081 Final Project Report
Do Men and Women Speak Differently About Monetary Policy?

Jiani Zhao

Abstract

This paper studies whether male and female central bankers differ in their monetary policy communication. Using English-language central bank speeches since 2015, it combines dictionary measures, supervised text classification, and word-embedding methods to analyze gendered language patterns. Dictionary-based hawkish–dovish scores show substantial overlap between male and female speakers, with little difference in average policy stance. In contrast, supervised classifiers predict speaker gender well above chance, suggesting that gender differences extend beyond simple tone. Word-embedding analyses further show that key policy terms, such as inflation and growth, are embedded in different semantic contexts across genders. Overall, the results indicate that gender differences in central bank communication reflect variation in framing and emphasis rather than systematic differences in policy stance.

Keywords: Central bank communication, Gender, Text-as-data

1. Introduction

1.1. Motivation

Central bank communication is a core policy tool in modern monetary policy-making, shaping market expectations, public understanding, and institutional credibility alongside formal interest rate decisions. As transparency has increased, the framing of policy messages has become increasingly important for how monetary policy is interpreted and trusted. At the same time, gender representation in central banking remains uneven, raising questions about whether diversity in leadership is reflected in communication practices. This project is motivated by the need to better understand whether male and female central bankers differ in how they communicate monetary policy, particularly with respect to tone and the framing of key economic concepts.

*Corresponding author
Email address: jz951@georgetown.edu (Jiani Zhao)

1.2. Previous Work

A large literature documents the informational content of central bank communication and its effects on expectations (Apel and Grimaldi, 2012). Sentiment extracted from central bank minutes has been shown to help forecast subsequent policy rate decisions, demonstrating that qualitative communication conveys economically meaningful information (Apel and Grimaldi, 2012). Central banks also use both interest rate decisions and language as complementary tools for managing expectations, particularly when conventional policy instruments are constrained (Hansen and McMahon, 2016).

More recent work applies automated text analysis to study the tonal and emotive dimensions of monetary policy communication. The tone of European Central Bank introductory statements contains predictive information about policy decisions (Baranowski et al., 2021). Dictionary-based sentiment measures can be informative, but results are sensitive to dictionary construction and word selection (Rutkowska and Szyszko, 2024).

While this literature establishes the importance of central bank communication, existing studies largely focus on aggregated sentiment or treat institutions as homogeneous communicators. To date, little attention has been paid to whether gender characteristics shape how monetary policy messages are framed. This study addresses this gap by examining whether male and female central bankers employ different linguistic styles and framing strategies when communicating about monetary policy.

1.3. Research Questions

This project investigates whether male and female central bankers communicate differently about monetary policy, both in tone and in semantic framing. Specifically, it asks:

1. Do male and female central bankers use systematically different language when discussing monetary policy decisions?
2. Are there gender differences in how hawkish and dovish concepts are framed and contextualized?
3. Do key monetary policy terms—such as inflation, growth, and stability—carry different semantic meanings for male versus female speakers?

1.4. Summary of the Report

To address these questions, the report analyzes English-language central bank speeches from the Central Bank Speeches (CBS) dataset using multiple text-as-data methods. First, supervised classification models are used to assess whether speaker gender can be predicted from full-text speeches, providing evidence of systematic linguistic differences. Second, dictionary-based hawkish–dovish measures are applied using established central bank lexicons to compare policy

tone across gender. Third, word-embedding models are estimated separately for male and female speech corpora to examine differences in the semantic neighborhoods of key monetary policy terms. By combining tone-based and semantic approaches, the report offers a comprehensive assessment of gender differences in central bank communication and contributes to the literature on monetary policy communication and gender representation in economic governance.

2. Data and Methods

2.1. Data Source and Sample Construction

The primary data for this study come from the Central Bank Speeches (CBS) dataset, an open-access corpus that compiles public speeches by central bank officials from more than 100 monetary authorities worldwide. The dataset includes the full text of speeches along with rich metadata such as speaker name, gender, institutional affiliation, country, date, and speech title. The CBS dataset is well suited for analyzing monetary policy communication because it provides standardized and systematically collected textual data across countries and time. The analysis focuses on speeches delivered in English or officially translated into English to ensure linguistic comparability across speakers. The sample is further restricted to speeches delivered between 2015 and 2024 in order to keep the dataset manageable and to focus on communication patterns in the most recent decade. To ensure sufficient representation of female speakers, the analysis includes only institutions and years in which both male and female central bankers appear with adequate frequency.

2.2. Unit of Observation and Variables of Interest

The unit of observation is an individual central bank speech. Each observation consists of the full speech text and associated metadata. The key independent variable of interest is speaker gender, coded as a binary indicator (female vs. male) based on the CBS dataset's gender classification. The primary textual outcome variables vary by method. For supervised classification, the outcome is speaker gender and the predictors are word-frequency features derived from the speech text. For dictionary-based analysis, the outcome of interest is a hawkish–dovish tone score, computed as the relative frequency of hawkish versus dovish terms using established central bank communication lexicons. For semantic analysis, the outcomes are word embeddings that capture the contextual meaning of key monetary policy terms such as inflation, growth, and stability.

2.3. Data Wrangling and Preprocessing

Several preprocessing steps are applied to prepare the textual data for analysis. First, speech texts are cleaned by removing punctuation, numbers, URLs, and non-informative symbols. All text is converted to lowercase to standardize word forms. Second, common English stopwords are removed to reduce noise, and words are stemmed to their root forms to consolidate similar terms (e.g., inflation and inflationary). Third, speeches with extremely short text or missing metadata are excluded to ensure analytical reliability. For methods that rely on word-frequency representations, a document–feature matrix is constructed using trimmed vocabularies that retain only reasonably frequent terms, reducing sparsity and computational burden. For semantic analysis, the corpus is split into male and female subsets, and word co-occurrence matrices are constructed using fixed context windows prior to estimating word embeddings. To ensure comparability across groups, vocabulary alignment and frequency thresholds are applied consistently across gender-specific corpora.

3. Analysis

This project uses multiple text-as-data methods to study gender differences in central bank communication. Because tone, framing, and meaning capture different aspects of language, the analysis combines supervised classification, dictionary-based sentiment measures, and word-embedding models. Each approach captures a distinct dimension of communication and helps validate the results from the others.

3.1. Supervised Text Classification

As a first step, the analysis tests whether speaker gender can be predicted from the full text of central bank speeches. Supervised classifiers using word-frequency features(Naive Bayes and regularized logistic regression) are trained on a subset of the data and evaluated on a held-out test set. Accuracy above chance indicates systematic linguistic differences across gender. This approach does not rely on predefined dictionaries, allowing patterns in language use to emerge directly from the data.

3.2. Dictionary-Based Hawkish–Dovish Tone Analysis

To assess differences in policy tone, the analysis uses dictionary-based measures of hawkish and dovish language drawn from the monetary policy literature. Each speech is assigned a hawkishness score based on the relative frequency of tightening- versus easing-related terms. This approach allows for direct comparison with prior studies and yields an interpretable measure of policy stance. Tone scores are then compared across gender. While transparent and widely used, dictionary methods cannot fully capture contextual meaning, motivating the use of semantic analysis.

3.3. Word Embeddings and Semantic Framing

To move beyond, the analysis uses word-embedding models to study how key monetary policy concepts are framed in context. Embeddings are trained separately on male andemale speech corpora to capture semantic relationships based on word co-occurrence. Comparing the semantic neighborhoods of terms such as inflation, growth, and stability reveals differences in meaning and emphasis that are not captured by word frequencies or dictionary-based measures. Together, these methods provide a complementary view of gendered communication in monetary policy. Classification models test whether language differs systematically by gender, dictionary-based measures capture differences in policy tone, and word embeddings uncover deeper variation in framing and meaning. Using multiple approaches reduces reliance on any single method and strengthens confidence in the results.

4. Results

This section focuses on the dictionary-based and semantic results, which form the main empirical findings of the study. Although supervised classification confirms that language differs systematically by gender, the analysis centers on policy tone and semantic framing, as these approaches are more interpretable and directly relevant for understanding monetary policy communication.

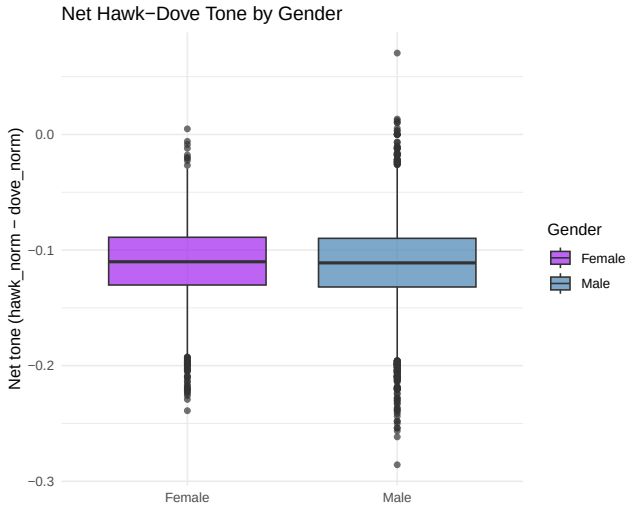

Figure 1: Net hawk–dove tone by speaker gender.

Figure 1 shows the distribution of net hawk–dove tone by gender. The two distributions overlap substantially, and the median net tone is very similar for

male and female speakers, indicating that average policy tone does not differ markedly by gender. In both groups, speeches tend to be net dovish on average, as reflected by negative median values. While central tendencies are nearly identical, male speakers exhibit slightly greater dispersion, with more observations in both the upper and lower tails of the distribution. Female speakers, by contrast, show a more concentrated distribution with fewer extreme values. Overall, these results suggest that gender differences in monetary policy communication are not primarily driven by average tone, but rather by differences in variability and extremity of language use. This motivates a closer examination of tone across topics and semantic contexts in subsequent analyses.

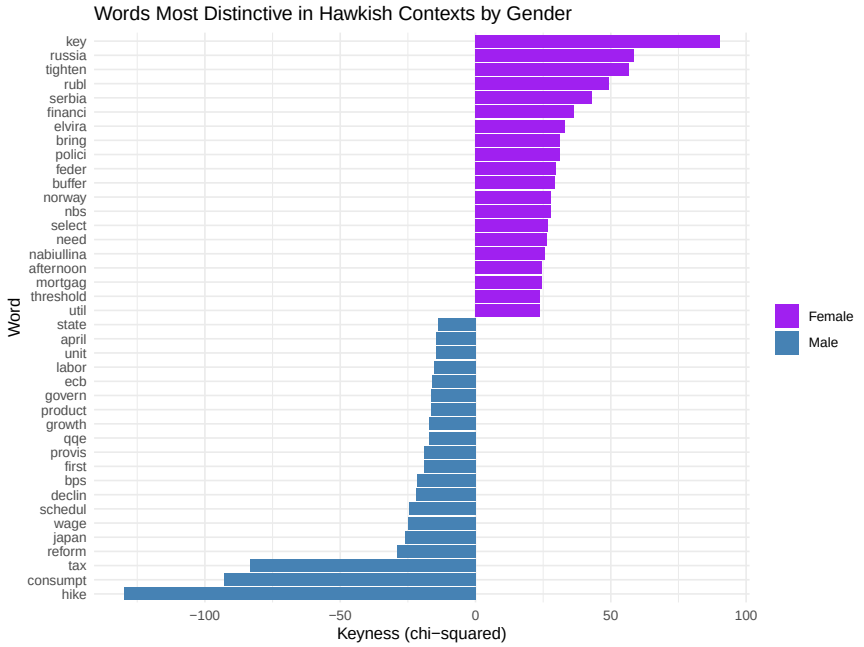

Figure 2: Words most distinctive in hawkish contexts

Figure 2 and Figure 3 present keyness analyses of the words most distinctive in hawkish and dovish monetary policy contexts by speaker gender. The figures compare gender-associated lexical patterns within dictionary-defined policy stances, highlighting differences in how male and female central bankers frame policy discussions rather than differences in overall tone.

In Figure 2, Female speakers' dovish language is dominated by institutional references and contextual framing, including terms such as feder, fomc, reserve, and country- or leader-specific references. Their language also includes household- and finance-related terms such as mortgage, payment, and consumption. In contrast, male speakers' dovish contexts are more strongly associated with macroeconomic aggregates and policy tools, including stimulus, monetary, reform, and

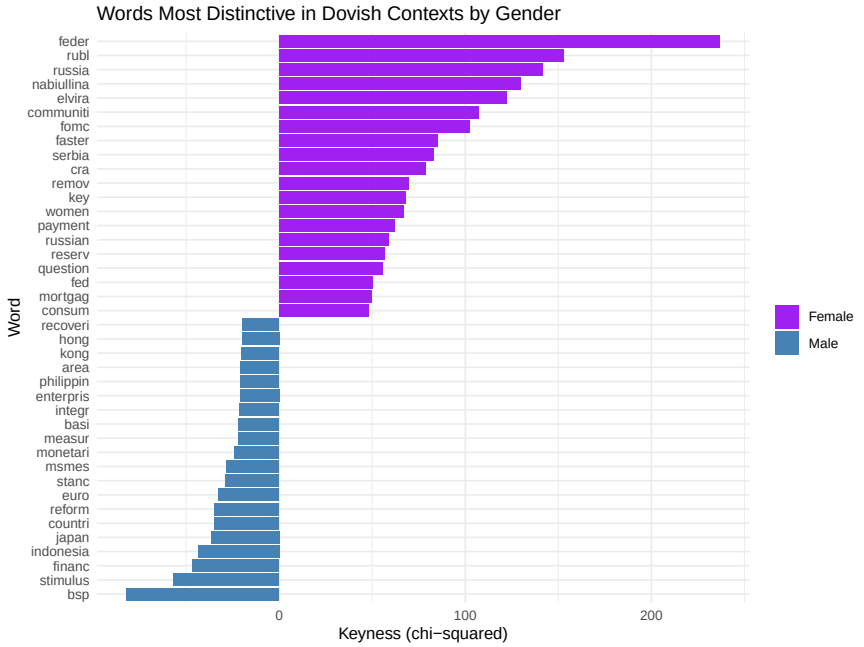

Figure 3: Words most distinctive in dovish contexts

growth.

In Figure 3, Female speakers' hawkish language emphasizes explicit policy signaling and institutional language, with terms such as tighten, policy, buffer, and threshold. Male speakers' hawkish contexts, by contrast, feature more concrete policy actions and outcomes, including hike, bps, tax, wage, and consumption. Together, these figures show that gender differences in monetary policy communication emerge most clearly in how policy stances are framed, even when overall tone is similar.

Figure 4 compares average net policy tone across topics by gender.

Across all topics, both male and female speakers exhibit negative average net tone, indicating that communication is generally dovish regardless of subject matter. Within each topic, gender differences in mean tone are small and highly overlapping. No single topic drives systematic differences in overall tone across gender. These results suggest that gender differences in monetary policy communication are not primarily driven by topic selection or topic-specific tone, but instead reflect differences in how similar topics are discussed.

Altogether, these results reinforce the finding that gender differences in communication are modest in terms of average tone. The clearer differences observed in the keyness and semantic analyses therefore reflect differences in framing and vocabulary within topics, rather than differences in the policy stance itself.

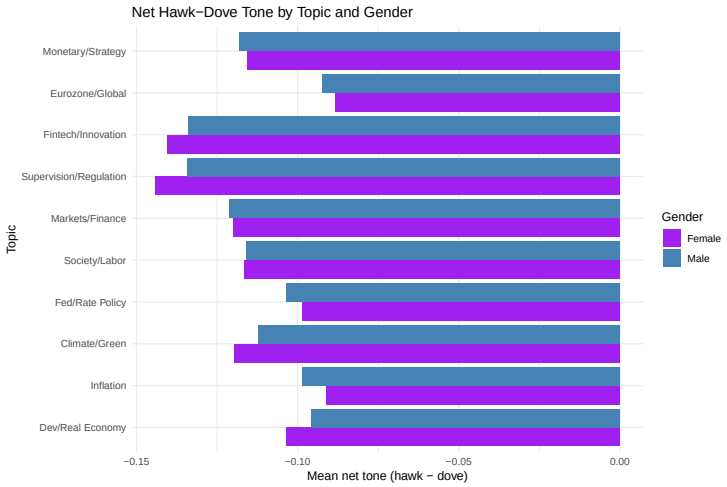

Figure 4: Net hawk–dove tone by topic and gender

To examine whether these differences reflect deeper shifts in meaning rather than surface-level vocabulary, the embedding analysis uses "tighten" as a focal term. Tightening is a core monetary policy concept that appears frequently across institutions and topics, making it well suited for comparing semantic associations across speakers.

Figure 5 compares the semantic similarity of words to "tighten" in female and male embedding spaces. Each point represents a neighboring word, with the dashed 45-degree line indicating equal similarity across genders. Several words lie close to this line, suggesting shared semantic associations with tightening across male and female speakers.

However, notable deviations from the diagonal reveal systematic differences in framing. Words such as pressure, gradual, anticipate, and slow are more strongly associated with tighten in male embeddings, while terms related to policy stance and instruments, such as policy, rate, and stance, have higher similarity in female embeddings. These patterns indicate that, even when discussing the same tightening concept, male speakers tend to frame tightening in terms of economic pressures and adjustment dynamics, whereas female speakers more often anchor tightening to formal policy actions and instruments. This result reinforces the conclusion that gender differences in monetary policy communication arise primarily through differences in semantic framing rather than differences in overall tone.

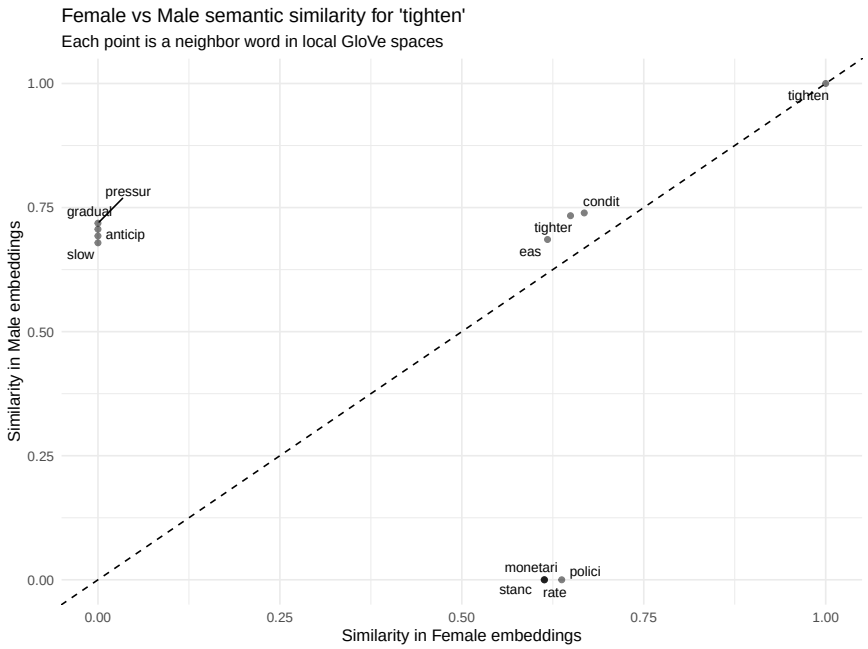

Figure 5: Female vs Male semantic similarity for 'tighten'

5. Discussion

This project finds clear evidence that male and female central bankers communicate differently about monetary policy. Across methods, speaker gender can be predicted from full-text speeches at rates above chance, indicating systematic differences in language use. Dictionary-based results point to modest but meaningful differences in policy tone, with male speakers more likely to use tightening-oriented and rule-based language, and female speakers more likely to emphasize labor markets, households, and broader economic conditions.

Most importantly, word-embedding analysis shows that key concepts such as inflation, growth, and stability are framed differently across gender, reflecting differences in contextual meaning rather than simple word frequency. Together, these results suggest that gendered differences in central bank communication extend beyond tone to deeper patterns of semantic framing.

5.1. Contributions & Future Directions

This study contributes to the literature on monetary policy communication and gender in several ways. It moves beyond existing work that focuses on policy topics or aggregate tone by examining how policymakers frame and contextualize

core economic concepts. It also provides systematic, cross-institutional evidence of gendered differences in monetary policy language.

Methodologically, the project shows the value of combining classification models, dictionary-based tone measures, and word embeddings to study policy communication. Taken together, the findings suggest that diversity in central banking may shape not only policy decisions but also the way policy messages are communicated and interpreted.

There are several natural directions for future work.

One extension would be to link linguistic differences more directly to outcomes such as asset price movements or changes in inflation expectations, to assess whether differences in communication style have measurable effects on audiences. The analysis could also be extended by incorporating richer contextual controls, such as macroeconomic conditions, crisis periods, or institution, specific communication norms to better isolate the role of gender.

On the methods side, embedding regression or dynamic embeddings could be used to study how the meaning of key monetary terms evolves over time and across institutions. In addition, scaling methods such as Wordfish could be applied to estimate latent policy positions directly from text and compare their distribution across speakers.

Finally, future research could move beyond binary gender classifications or examine how communication varies by role, seniority or institutional culture. Together, these extensions would deepen understanding of how diversity in central banking shapes both policy communication and its economic consequences.

References

Apel, M., Grimaldi, M.B., 2012. The Information Content of Central Bank Minutes. Technical Report 261. Sveriges Riksbank. doi:10.2139/ssrn.2092575. sveriges Riksbank Working Paper Series.

Baranowski, P., Bennani, H., Doryń, W., 2021. Do the ecb's introductory statements help predict monetary policy? evidence from a tone analysis. European Journal of Political Economy 66, 101967. doi:10.1016/j.ejpoleco.2020.101967.

Hansen, S., McMahon, M., 2016. Words versus deeds: How central banks manage expectations. URL: https://cepr.org/voxeu/columns/words-versus-deeds-how-central-banks-manage-expectations. voxEU.org.

Rutkowska, A., Szyszko, M., 2024. Dictionary-based sentiment analysis of monetary policy communication: On the applicability of lexicons. Quality & Quantity 58, 5421–5444. doi:10.1007/s11135-024-01896-9.