The Language of Central Banking: Measuring Complexity and Market Impact

Alex Ye

Senior Essay, Yale Department of Economics*

April 2025

Abstract

This thesis examines the linguistic complexity of monetary policy communications from the Federal Reserve, Bank of England, and European Central Bank. It finds that, of the three, FOMC statements tend to be the most semantically complex—using difficult grammatical structures and longer sentences—while Bank of England statements tend to be the most conceptually complex, relying on technical jargon and abstract economic topics. It also investigates whether complexity affects market participants' understanding of policy by examining the relationship between language complexity and market expectations of forward-looking volatility. The Euro area presents the strongest evidence that language complexity matters: greater semantic complexity is associated with higher expectations of future volatility, while greater conceptual complexity appears to lower these expectations, suggesting a clearer understanding of monetary policy.

* I would like to thank my advisor Bill English for his invaluable mentorship, thoughtful advice, and continuous encouragement. I also want to thank my friends and family for their constant support during my academic journey. All errors are my own.

1 Introduction

The Federal Reserve ("Fed") sets monetary policy in response to economic developments with the goal of maintaining maximum employment and stable prices. Although the Fed primarily achieves this objective through the adjustment of the policy rate, it is well documented that the manner in which a central bank communicates can have a significant effect on the economy as well (Woodford 2005, Blinder et al. 2008). Effective communication by central banks can help improve households' understanding of the state of the economy and guide their perception of the future path of policy. Moreover, effective communication is necessary in maintaining the public's trust in the monetary authority (Macklem and Vardy 2023).

Central bank communications have evolved significantly in recent years, driven in part by the adoption of unconventional monetary policies such as quantitative easing in response to the Global Financial Crisis (GFC) and the COVID-19 pandemic. During the GFC, the Fed increasingly relied on forward guidance as a tool, which could only be effective if the public understood and believed its objectives (Yellen 2012). In such cases, explaining policy to the public may require the central bank to utilize economic language that may be difficult for the the ordinary layperson to understand.

This thesis examines the monetary policy communications of three major central banks: the Federal Reserve System (Fed), the Bank of England (BoE), and the European Central Bank (ECB). The goals are twofold. The first is to quantify the complexity of monetary policy communication and examine how it has evolved over time. The analysis focuses on statements given by the central banks on the days that the policy rate is set, as well as central bank meeting minutes. For each document in the sample, three measures of language complexity are calculated: the Flesch-Kincaid score, the Dale-Chall score, and the Conceptual Complexity Index (CCI). The Flesch-Kincaid and Dale-Chall are *semantic* measures, assessing syllables and sentence length; on the other hand, the CCI is a *conceptual* measure, assessing the extent of economic jargon and breadth of topics covered (Kincaid et al. 1975; McMahon and Naylor 2023). The second objective is to investigate the extent to which the complexity of monetary policy communication affects market participants' reaction to monetary policy. Specifically, the relationship between statement complexity and market-implied volatility is investigated. Changes in forward-looking volatility, measured through financial instruments such as the VIX, are assessed on central bank announcement days. If it is the case that market participants' have a weaker understanding of the path of monetary policy following a complex statement, then one might observe heightened expectations of market volatility going forward.

Ultimately, FOMC statements are found to be the most semantically complex out of the three central banks, and BoE statements are the most conceptually complex. Results for the relationship between language complexity and forward-looking expected volatility are mixed; the relationships are only significant for the ECB, where an increase in the Flesch-Kincaid score tends to increase forward-looking volatility and an increase in the CCI tends to decrease forward-looking volatility, at 1% and 5% significance levels respectively. Notably, the results are economically meaningful, with a one standard deviation increase in the Flesch-Kincaid score causing the European implied volatility index (VSTOXX) to rise ∼1.9% and a one standard deviation increase in CCI causing VSTOXX to decline by ∼1.0% on the announcement day. These results suggest that higher semantic complexity may have some effect of reducing understanding of the statement, whereas higher conceptual complexity may help improve understanding, perhaps by providing a more complete understanding of the policy decision.

The rest of this thesis is organized as follows. Section 2 describes related literature as well as hypotheses that have been presented on the topic of central bank communication. Section 3 provides an overview of monetary policy communication in the three central banks examined. Section 4 describes the data and methodology. Section 5 presents and discusses results. Section 6 concludes.

2 Literature Review

2.1 Analyzing Complexity in Central Bank Statements

Over the past century, substantial research has been done on the subject of readability, with a particular focus on identifying the factors that influence a reader's ability to comprehend and retain material. One of the earliest and most influential contributions to this field came from Rudolf Flesch. In 1943, Flesch's doctoral dissertation in educational research introduced a statistical formula designed to objectively measure readability (Flesch 1948). The formula assessed readability based on three key factors: average sentence length, the number of affixes, and references to people. Independent studies later validated Flesch's formula, finding that it aligned closely with student and teacher ratings of psychology textbooks. Additionally, the studies revealed that advertisements and articles rated as "more readable" by the formula tended to experience higher readership.

World War II would highlight the need for effective communication with large audiences, fuelling further advancements in readability research. In 1948, Rudolf Flesch refined his original readability formula by incorporating additional factors, such as syllables per word and the frequency of "personal" words, believing these elements would make texts more engaging. That same year, Dale and Chall (1948) developed their own readability formula using a word list of 769 words that 80 percent of fourth-graders recognized. The resulting score demonstrated over 90 percent correlation with assessments made by readability experts, as well as evaluations by children and adults.

Kincaid et al. (1975) further extended Flesch's work by creating a modified score that measured the grade level required to comprehend a text. This adaptation was designed to meet the practical needs of the U.S. Navy. Later, Chall and Dale (1995) revisited the original Dale-Chall formula, modifying it to use a more extensive list of 3,000 familiar words for improved accuracy.

This thesis is most closely related to the work done by Smales and Apergis (2017), who utilize measures of linguistic complexity such as word count and readability scores to gauge the effects of complex statements on financial markets. They find that both word count and readability are statistically significant predictors of the quoted spread, quoted depth, and number of quotes of the 10-year Treasury note, after controlling for surprises in the target rate and path of policy. The authors conclude that "a long statement which is easy to understand may help to increase bond market liquidity in the period following the FOMC announcement", consistent with the hypothesis that more complex information leads to differing interpretations that decrease liquidity and increase volatility. My thesis builds on these results by analyzing communication of the ECB and BoE in addition to the FOMC, along with estimating a QE surprise factor in addition to the target and path surprises, as in Smales and Apergis (2017).

The specific score used by Smales and Apergis (2017) is the Flesch-Kincaid readability level, which is lower for statements with longer words and for statements with longer sentences. Bulir et al. (2014) conduct a similar analysis using the Flesch-Kincaid grade level score. Using inflation reports from the Czech National Bank, the European Central Bank, the Bank of England, and Sveriges Riksbank, they identify a negative relationship between the clarity of a report and asset return volatility during the period prior to the Great Financial Crisis. This relationship, however, disappears during the financial crisis. Their finding suggests that the benefits of clearer communication may not be stable across different economic environments.

While most research focuses on the effects of clear communication on financial market participants, McMahon and Naylor (2023) conduct an experiment that attempts to evaluate the general public's understanding of central bank statements. In addition to the traditional semantic measure of complexity provided by Flesch-Kincaid, they also construct a 'Conceptual Complexity Index' that is increasing in the proportion of jargon terms used, the breadth of distinct jargon terms within a given topic, and the number of different topics covered. The idea is that while semantic measures may penalize documents for having longer words and longer sentences, the conceptual measure may do a better job of assessing how difficult it is for an ordinary person to understand the document. In an experiment involving 1,859 participants, they find that higher conceptual complexity reduced respondents' perceived understanding of the report, their actual understanding of the report, and their sentiment (trust) towards the central bank.

A number of papers have argued for the benefits of clear communication. Woodford (2005) contends that "the increased willingness of... Alan Greenspan to speak openly about both current policy decisions and the Committee's view of likely future policy has greatly increased the ability of markets to anticipate Fed policy." He argues that greater predictability improves the central bank's ability to achieve its stabilization objectives. Blinder (2008) surveys the literature on central bank communication, concluding that while "no consensus" has yet emerged on what constitutes "best [practices]", communication nevertheless is an important tool that has the ability to move financial markets and improve the predictability of monetary policy decisions. Specifically, Blinder notes that "expectations about future central bank behavior provide the essential link between short rates and long rates", emphasizing that effective communication strengthens the central bank's influence on the economy's trajectory.

Amato, Morris, and Shin (2002) highlight three key drivers behind the heightened emphasis on central bank communication: greater independence prompting demands for accountability, the implementation of inflation targeting, and the rising impact of financial markets and expectations. They caution that market participants may overreact to information released by the central bank, particularly when obtaining private information is costly. This could suppress the flow of private information, creating risks if public information is incorrect.

2.2 Assessing Asset Price Reactions to Central Bank Statements

This thesis also follows from a second strand of literature that attempts to measure financial market reactions to monetary policy announcements. Specifically, these studies analyze high-frequency price movements following monetary policy decisions to evaluate the effects of monetary policy on asset prices. Kuttner (2001) measures daily changes around the current-month or next-month federal funds rate to estimate the surprise component of the announcement, as well as the effects of policy rate changes on Treasury yields. Gurkaynak, Sack, and Swanson (2005), henceforth GSS, show that the surprise component of an announcement can be decomposed into different factors by analyzing changes in both short-term rates and long-term yields following monetary policy announcements. They identify two explanatory factors: a "current federal funds rate target" factor and a "future path of policy" factor. The factors are derived by estimating two principal components from a matrix of intraday price movements for various Fed announcements, then rotating the principal components so that they can be structurally interpreted as a "target" factor and a "path" factor. The authors then measure the effects of the surprise factors on bond yields and stock prices, finding significant but differing effects.

Swanson (2017) extends the methodology of GSS to separately identify the effects of large-scale asset purchases (LSAPs), in addition to changes in the target and path factors. He associates the path factor with forward guidance, finding that both the forward guidance and LSAP factor have statistically significant effects on financial markets, comparable in magnitude to the target rate effect before 2009. Altavilla et al. (2019) applies the methods of GSS and Swanson (2017) to the Euro area, using two intraday windows to account for both the ECB press release and press conference. They identify four factors: "Target", "Timing", "Forward Guidance", and "QE". The "Timing" factor is similar to "Forward Guidance", but reflects changes in expectations over a shorter time horizon of around six months, while the "Forward

Guidance" factor captures changes over a two-year horizon. Their findings show that "Target" surprises have effects only on the short end of the yield curve while "Timing", "Forward Guidance", and "QE" all influence longer term rates, albeit in different ways. Similarly, Braun et al. (2023) uses the methodology of Swanson (2017) to identify the different dimensions of UK monetary policy, finding that QE has a statistically significant effect on the economy when the policy rate is at the effective lower bound.

Nakamura and Steinsson (2018) find evidence of a "Fed information effect" during monetary policy surprises. Using a high-frequency identification approach to estimate the causal effects of monetary policy surprises, they find that in contrast to standard macroeconomic models, monetary tightening shocks raise expectations for output growth. They develop a model for the Fed information effect, in which market participants interpret part of an unexpected change in the target rate as information about the neutral rate of interest. Their analysis suggests that the information content of a monetary surprise is a key factor in determining its causal effects.

More recently, Swanson and Jayawickrema (2024) have applied the methodologies of Kuttner (2001) and GSS to other types of Fed communication beyond announcements, such as post-meeting press conferences, speeches, Congressional testimony, and FOMC meeting minutes release. They find that speeches by the Fed chair have greater effects on financial markets compared to other types of communication, while minutes releases are less significant. They also note that post-FOMC press conferences have become more important over time.

3 Overview on monetary policy communication

3.1 Federal Reserve

The modern-day Federal Reserve communicates frequently with the public. After each of its eight regularly-scheduled annual FOMC meetings, the Fed issues a statement and holds a press conference. Three weeks later, it releases the meeting minutes. At every other meeting, the Fed also publishes a Summary of Economic Projections (SEP), which includes FOMC participants' forecasts for GDP growth, inflation, unemployment, and interest rates. Additionally, policymakers often deliver speeches and participate in interviews to inform the public about the Fed's current goals.

The Fed's practice of communicating often is a relatively recent development. Prior to the 1990s, central bank operations were shrouded in mystery, and the FOMC would not reveal its interest rate decisions. Alan Greenspan noted in 1993 that "The Federal Reserve has a reputation, along with other central banks, of being secretive" (Greenspan 1993). In February 1994, the Fed began issuing post-meeting statements, although only after meetings where the federal funds rate was changed. In May 1999, the Fed went further and started releasing statements after every meeting, regardless of whether the policy rate was changed. During this period, the Fed also started publishing FOMC meeting minutes "in their current form", beginning in 1993 (Jefferson 2024). With the December 2004 meeting, the Fed moved from releasing meeting minutes three days following the subsequent meeting to three weeks after the current meeting, providing market participants with quicker access to the Fed's discussions (Danker and Luecke 2005).

The trend towards greater central bank transparency gained further momentum in the 21st century. In 2007, the Fed introduced the SEP, and in 2011, Chairman Bernanke began holding press conferences after every other meeting. By 2019, Chairman Powell had expanded this practice, holding press conferences following every FOMC meeting.

What motivated the push towards greater communication and transparency? Fed Historian Jonathan Rose notes in his 2024 article "Transparency" that central bankers today view transparency as critical in maintaining central bank independence, particularly in the face of heightened scrutiny from both legislators and the general public. He also notes that the increase in communication has coincided with an increase in government transparency generally. Moreover, communication is increasingly viewed as a tool that strengthens the effectiveness of monetary policy. Woodford (2005) contends that increased communication not only reduces market participants' uncertainty but also allows the Fed to better achieve its desired economic outcomes by aligning expectations.

In a 2012 speech describing the evolution of Fed communications, former Fed Chair Janet Yellen highlights both public pressure and academic developments as catalysts for increased transparency. She notes that the "[secretiveness prior to the 1990s] regarding monetary policy decisions clashed with the openness regarding government decisions expected in a democracy, especially since Federal Reserve decisions influence the lives of every American."

Public demand for greater transparency was evident as early as 1975, when a citizen sued the FOMC, calling for increased openness. During the same period, research published by Robert Lucas and others would challenge the traditional view that secrecy regarding monetary policy was the best approach. Since the public's expectations about future policy influence their spending and investment decisions today, monetary policy's impact depends not only on the current target rate but also on the anticipated path of rates going forward—a path that can only be understood through transparent communication.

These insights became even more critical during the 2008 Global Financial Crisis, when the Fed began to rely heavily on forward guidance about both the path of rates and large-scale asset purchases. Yellen explains that in order for forward guidance to have maximum effect, it "must be understood and believed by the public", something that is only possible if the public understands the Fed's objective and intentions.

3.2 European Central Bank

Like the Federal Reserve, the European Central Bank currently communicates often. The ECB releases a policy statement and holds a press conference after each Governing Council (GC) meeting about monetary policy, which occur every 6 weeks. Minutes are released about a month after the GC meeting, and staff economic forecasts are also released 4 times per year. GC members regularly give speeches, and the President and Vice President also give testimony in Parliament.

At its start in 1998, the ECB displayed greater transparency than the Fed, with the President making public statements and holding a press conference following GC meetings. Recall that the Fed did not hold press conferences after every meeting until 2019. With regard to meeting minutes, however, the ECB would only begin publishing the "accounts" of GC meetings in 2015. Notably, these accounts do not include individual statements from GC members nor voting results on policy decisions (de Haan and Hoogduin 2024).

The ECB faces a number of challenges that arise from its position as a supranational bank. Unlike the Fed, the ECB must communicate its message in over 20 different languages. Diverse audiences across the 20 member states exhibit varying levels of financial literacy and trust in the ECB, making it difficult to craft a message that gets through to every individual. Gardt et al.

(2021) write in the ECB Economic Bulletin that “the complexity of central bank communication and low levels of financial literacy, among other factors, can make it too ‘costly’ for people to pay attention to what central banks say.”

3.3 Bank of England

“Never explain, never excuse” was a favorite saying of Montagu Norman, the head of the Bank of England from 1920 to 1944. However, like the Fed and the ECB, the Bank of England has gradually increased communication with the public over time. In February 1993, the Bank began publishing the quarterly Inflation Report, with the goal of informing the public about inflationary trends and pressures. Specifically, the original Inflation Report covered recent price developments, recent cost developments, the outlook for inflation, and a conclusion. Notably, it initially provided little commentary on the outlook for policy, though this would change over time. The Inflation Report is now called the Monetary Policy Report and continues to be published quarterly to this day (Millard 2022).

The Bank of England Act of 1998 would further increase communication and transparency. The Act required that the Monetary Policy Committee (MPC) meet at least monthly, publish its decisions promptly, and publish meeting minutes within six weeks. Specifically, the Act required that the minutes contain the votes of each member. (The monthly meeting requirement would later be modified under the Bank of England and Financial Services Act of 2016. Under the recommendation of Fed governor Kevin Warsh, the MPC would be required to be meet only eight times per year.) Today, the Bank of England publishes the MPC’s decision along with the meeting minutes at 12 noon on the day of the announcement. At the four meetings where the quarterly Monetary Policy Report is published, the BoE governor also holds a press conference after announcing the policy decision.

4 Methodology

4.1 Processing Textual Data

The textual dataset analyzed includes statements and minutes from the FOMC, ECB, and BoE. Statements are selected because they serve as the primary channel for communicating policy actions to the public and are closely examined by market participants, enabling the analysis of asset price reactions in short windows around their release. Minutes are also included, as they provide valuable insight into the Committee's monetary policy outlook and are also closely scrutinized by market participants. While other forms of communication, such as press conferences and speeches, are also analyzable, statements and minutes are chosen due to their consistent format—which facilitates comparison over time—and their status as the official record of monetary policy decisions. Additionally, language complexity scores may be affected by filler words and less structured language used in impromptu texts like speeches and answers to questions, which is why statements and minutes are preferable for this analysis.

For the FOMC, my dataset includes all statements and minutes from federalreserve.gov for the period 2001-2024. For the ECB, the Monetary Policy Decision serves as the equivalent of a statement but offered little information beyond the policy action until the past decade. Instead, the statement delivered by the ECB President at the start of the press conference is used. I also gather minutes for the ECB, sourcing both statements and minutes from ecb.europa.eu. While statements are available for the entirety of 2001-2024, the ECB only began releasing minutes in 2015.

The Bank of England is unique in offering three forms of communication around policy meetings. The Monetary Policy Report (formerly the Inflation Report) provides the most comprehensive analysis of policy decisions and economic conditions. The Monetary Policy Summary (MPS) is a more concise version, tailored to a technical audience, while the Visual Summary, introduced in 2017, simplifies communication using charts and figures for a broader audience. For the purposes of my analysis, the MPS is most comparable to the FOMC statement and the ECB press statement. However, before August 2015, the MPS contained little information beyond the policy decision. Therefore, my dataset contains MPS from August 2015 to 2024, along with the minutes of the Bank of England's Monetary Policy Committee meetings from 2001 to 2024, sourced from bankofengland.co.uk.

The texts are downloaded as PDFs. They are then read using the open-source PyPDF2 package in Python. The texts are then cleaned as follows:

1. First, any content at the beginning or end of the PDF that is unrelated to the main body of the statement or minutes is removed. This includes elements such as the document

title, date, and time—typically found at the start—as well as information about media inquiries, which usually appears at the end.

2. Long lists of names listed in sections detailing voting decisions are removed. Since names can influence the complexity scores we calculate without directly affecting a reader's comprehension, their removal ensures a more accurate assessment.

3. Economic projections containing numerous numerical values are excluded, as numbers tend to lower the complexity score without necessarily making the text easier to understand.

In total, the dataset contains 191 FOMC statements, 191 FOMC meeting minutes, 80 BoE statements, 255 BoE minutes, 242 ECB statements, and 82 ECB meeting minutes.

4.2 Quantifying Complexity

After cleaning the PDFs, three complexity scores are calculated for each document: the Flesch-Kincaid Grade Level, the Dale-Chall Score, and the Conceptual Complexity Index. Because the scores were originally derived and applied on shorter texts, and because calculating the scores for large texts is computationally expensive, I instead take three randomly-selected 500-word samples from each text. The scores are then calculated for each 500-word sample, and then the final score for the document is taken to be the average score across all three samples.

The first score is the Flesch-Kincaid grade level score. Developed in 1975 by J. Peter Kincaid for the U.S. Navy, the score is calculated according to the following formula:

$$FK = 0.39\frac{\text{total words}}{\text{total sentences}} + 11.8\frac{\text{total syllables}}{\text{total words}} - 15.59$$

The output of the Flesch-Kincaid formula is a number that roughly corresponds with the years of education necessary to understand the text. The score is increasing as both individual words and sentences become longer, making it a proxy for complexity based purely on semantics.

The second score is the Dale-Chall Score. Originally developed in 1948, the score is higher for texts that contain a greater number of words that fall outside a list of the most common English words. The score was updated in 1995, expanding the word list to 3,000 common English words.

The Dale-Chall score is calculated according to the following formula:

$$DC = 0.1579\left(\frac{\text{difficult words}}{\text{words}} \times 100\right) + 0.0496\left(\frac{\text{words}}{\text{sentences}}\right)$$

Here, 'difficult words' is the number of words that fall outside the 3,000 word list. Scores of 4.9 or lower correspond to texts that are easily understood by a 4th-grade student, 5.0-5.9 corresponds to 5th- or 6th-grade student, and so on, with 9.0 and above corresponding to a college student.

The third score is the Conceptual Complexity Index (CCI). Developed by Michael McMahon and Matthew Naylor (2023), the CCI measures how complex a text based on the *concepts* it contains—the more specialized and technical the concepts, the higher the score. Specifically, the CCI increases when:

1. Jargon words represent a higher proportion of the text.

2. Jargon words used in the text are spread across multiple different topics.

3. The text covers a wider range of topics.

McMahon and Naylor (2023) identify jargon words by creating a jargon dictionary containing 350 terms categorized into 10 different topics; these terms are identified based on lists of economic and financial terms published by the Economist, the Guardian, and Investopedia. Additional detail about the mathematical derivation of the CCI is provided in Appendix C.

After computing the CCI, they conduct an experimental study with 1,859 representative members of the UK public to test how different dimensions of complexity affect effective communication. Participants were randomly assigned to read reports with varying levels of semantic and conceptual complexity, and their understanding and sentiments towards the central bank were assessed via survey. By regressing the variables measuring understanding and sentiment on different measures of complexity, the study found that conceptual complexity, as measured by the Conceptual Complexity Index (CCI), was the primary driver of both. Additionally, the study suggested a potential "goldilocks" effect, where medium levels of conceptual complexity led to better understanding, but higher complexity caused diminishing returns.

Overall, the Flesch-Kincaid and Dale-Chall provide *semantic* measures of language complexity, due to their focus on syllables and sentence length. The Conceptual Complexity Index provides a *conceptual* measure of language complexity due to its focus on the actual meanings of words.

Figure 1 displays a comparison of central bank statements and meeting minutes in terms of complexity, across the three different central banks and the three different metrics of complexity. In most cases, complexity of meeting minutes was similar to the complexity of statements. This is somewhat striking, as one might expect statements to be simpler due to their broader audience.

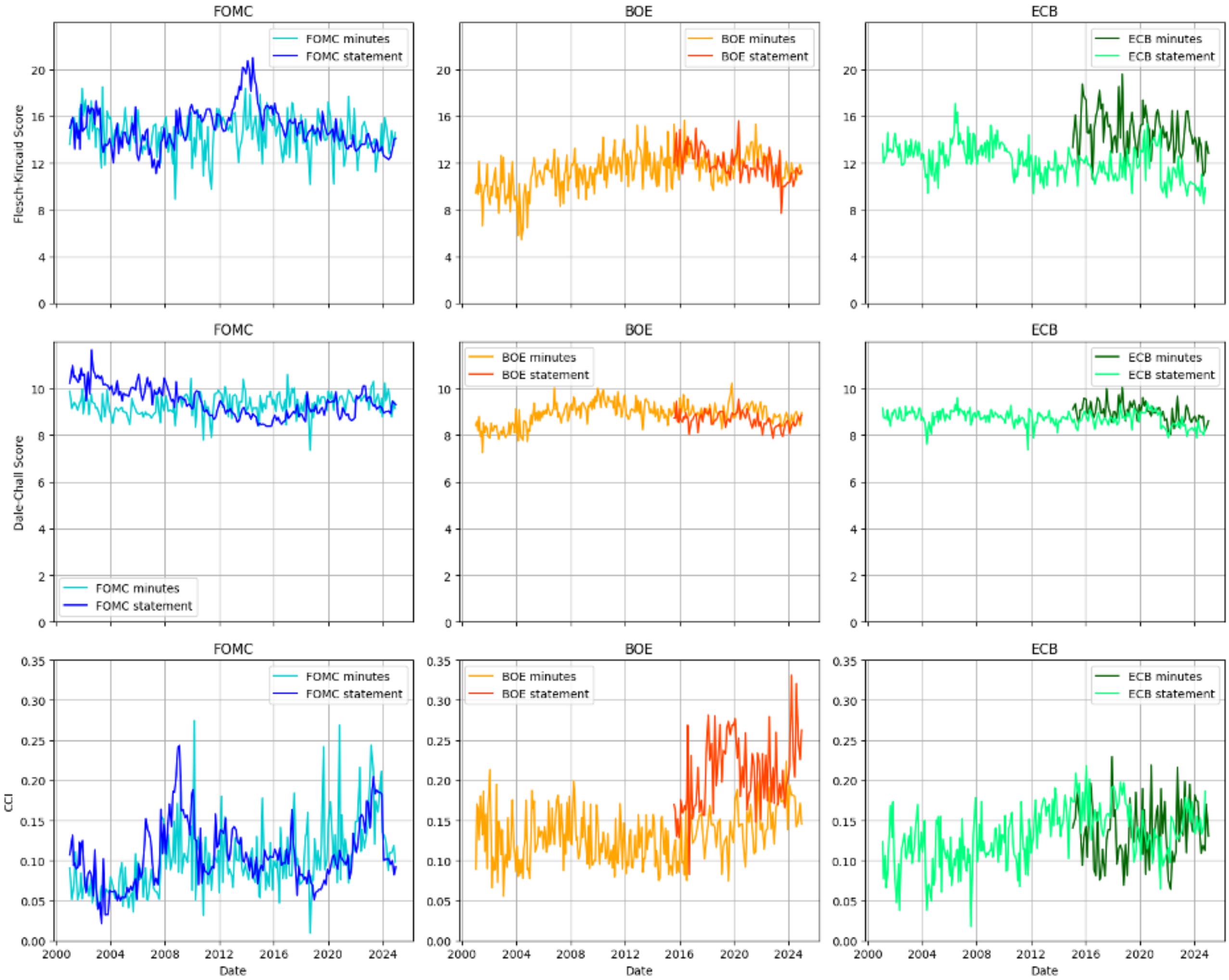

Figure 1: Statements and Minutes Comparison

For the FOMC, notable exceptions where the statement complexity diverged from the minutes included the GFC years of 2008-09 as measured by the CCI, as well as the post-2012 years as measured by the Flesch-Kincaid score. In both instances, statements became far more complex than minutes. However, the two instances reflected relatively different instances of policy communication. For instance, in 2008, communication exhibited a high level of conceptual complexity likely due to the novelty of measures used by the Fed to combat the crisis. These measures were wide-ranging, from bailouts to quantitative easing to special lending programs. The high dispersion of these topics would contribute to a higher conceptual complexity.

However, in 2012, the Fed was just beginning QE3, which coincided with a period of higher semantic complexity as measured by the Flesch-Kincaid score. Notably, there was no jump in conceptual complexity. A potential reason for this may have been that explaining QE programs made policymakers more prone to using lengthier sentences, but not necessarily using jargon from a wide range of topics.

For the BoE, semantic complexity as measured by Flesch-Kincaid is lower than that of the Fed on average, by about 3 grade levels. Notably, BoE minutes are similar to statements in complexity when measuring the Flesch-Kincaid or Dale-Chall score, but not when measuring the CCI. Instead, the CCI indicates that BoE statements are more conceptually complex than BoE minutes, which is somewhat surprising given that minutes describe internal discussions of central bankers, while statements are a press release geared towards a larger public audience.

ECB minutes and statements are most similar in complexity when measured by the Dale-Chall score, by diverge significantly for both the Flesch-Kincaid score and CCI. Minutes appear to be slightly more complex than statements across the semantic measures of Flesch-Kincaid and Dale-Chall, but about the same and slightly lower at times when measured by the CCI.

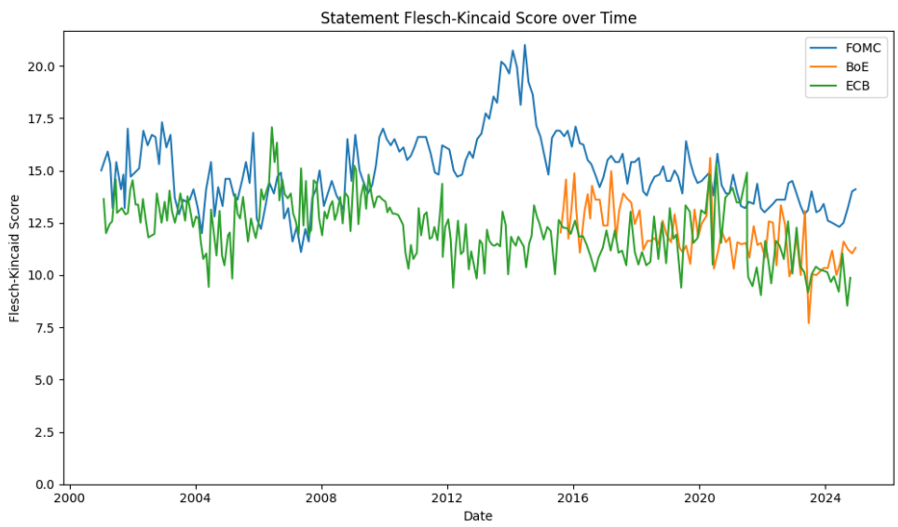

Figure 2: Flesch-Kincaid Score (Statement)

Figures 2 through 4 compare the statements across the different central banks, for all three complexity scores. Example excerpts from the most complex statements, as measured by the three scores, are provided in Appendix B.

Based on the Flesch-Kincaid score, the following observations are made:

1. FOMC statements appear to be the most complex, while ECB and BoE statements are simpler.

2. Following the 2008 Global Financial Crisis, complexity rose dramatically for FOMC statements, although the same cannot be said for ECB statements. During this time, the FOMC implemented drastic measures including quantitative easing to support the economy, whereas the ECB response was less aggressive.

3. In the past decade, complexity appears to have declined for all three central banks. This is consistent with central bank efforts to improve communication with the general public.

The average fluctuation in the Flesch-Kincaid score from meeting to meeting is 0.81 for the FOMC, 1.15 for the BoE, and 1.08 for the ECB. For a 500-word sample (∼20 sentences), a one-unit change in the Flesch-Kincaid score can be achieved by either adding or subtracting 51 syllables or by adding or subtracting 42 words, or by a combination of both. Specifically for FOMC statements, whose average word count is 350 (∼14 sentences), a fluctuation of 0.81 can result from a change of 24 syllables or 29 words.

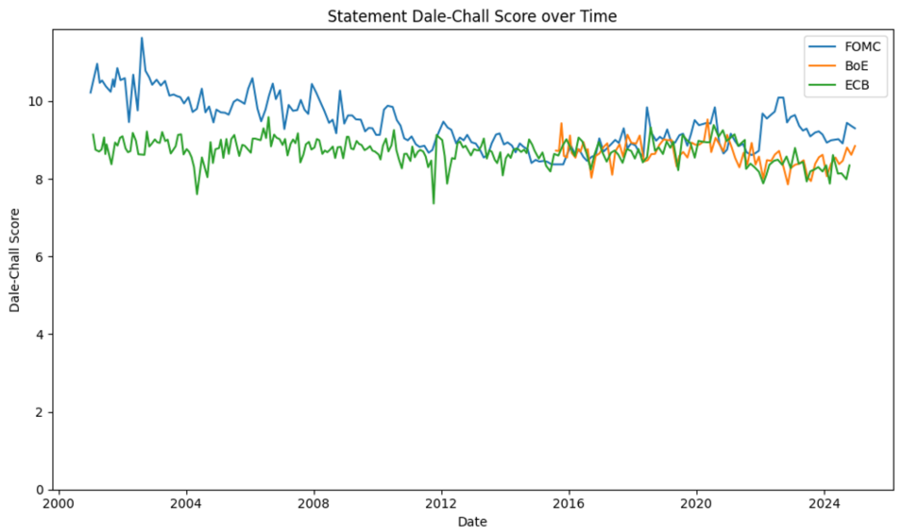

Figure 3: Dale-Chall Score (Statement)

Similar to Flesch-Kincaid, Dale-Chall is higher for texts with longer sentences. However, the key difference is that Dale-Chall identifies difficult words not by how many syllables but rather whether they fall outside a pre-defined word list. As a result, Dale-Chall accounts for words that have many syllables but are commonly used, making them easy to comprehend.

Based on Dale-Chall, the following observations are made:

1. As was the case with Flesch-Kincaid, FOMC statements appear to be the most complex, followed by BoE and ECB statements.

2. Both ECB and BoE statements have slightly declined in complexity in recent years, whereas it is less apparent if FOMC statement complexity has declined.

3. There was no post-2008 increase in complexity for the FOMC, which Flesch-Kincaid indicated.[1]

The average fluctuation in the Dale-Chall score from meeting to meeting is 0.24 for the FOMC, 0.29 for the BoE, and 0.25 for the ECB. For a 500-word sample, a 0.25-unit change in the Dale-Chall score can be achieved by either increasing the average sentence length by 5 words or raising the proportion of "difficult words" in the text by 1.58 percentage points, or by a combination of both.

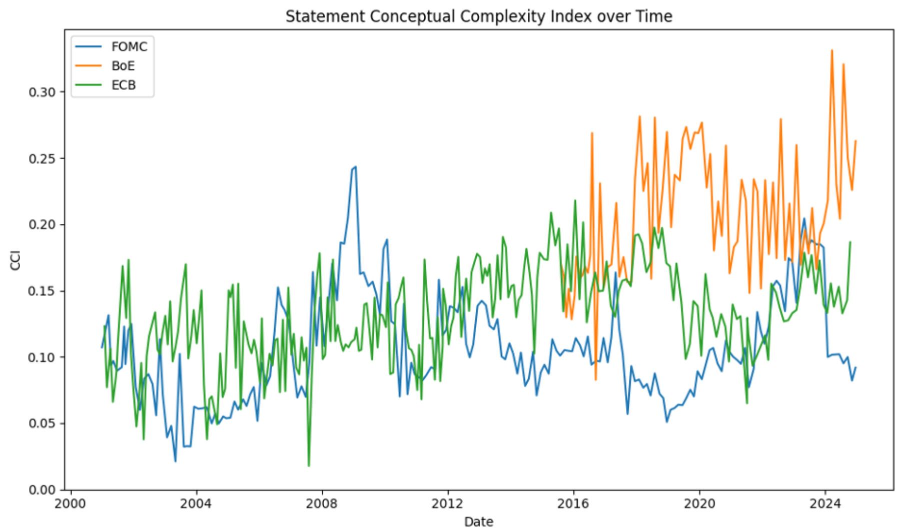

Figure 4: Conceptual Complexity Index (Statement)

Based on the CCI, the following observations are made:

1. FOMC statements now appear to be the least complex, followed by ECB. The BoE statements now appear to be the most complicated.

[1] See Appendix B for more detail.

2. FOMC statements increased sharply in CCI immediately following the 2008 Global Financial Crisis. ECB statements also increased in CCI somewhat during 2008, but exhibited a much greater increase in the mid-2010s, consistent with the ECB taking more aggressive and unconventional actions to support the economy during that time.

3. Both the FOMC and ECB statements exhibit an increase in CCI following the inflation of 2022, although the BoE statements do not exhibit this same increase.

The average fluctuation in the CCI from meeting to meeting is 1.8% for the FOMC, 4.7% for the BoE, and 2.7% for the ECB. For the CCI, the change required to generate an increase of approximately 5% varies depending on the number of jargon terms and number of topics. With 8 topics, a 5% increase in CCI can be achieved by adding jargon from a ninth topic or increasing the total jargon terms by 5%. Additional mathematical detail is provided in Appendix C.

Consistent with McMahon and Naylor (2023), BoE statements do not exhibit a trend-decline in CCI despite displaying some decline in the semantic complexity measures. Specifically, BoE statements exhibit higher levels of conceptual complexity (over 0.25) than in 2015 (closer to 0.20).

Finally, the minutes across the different central banks are compared, for all three scores. The comparison plots are displayed below in Figures 5-7.

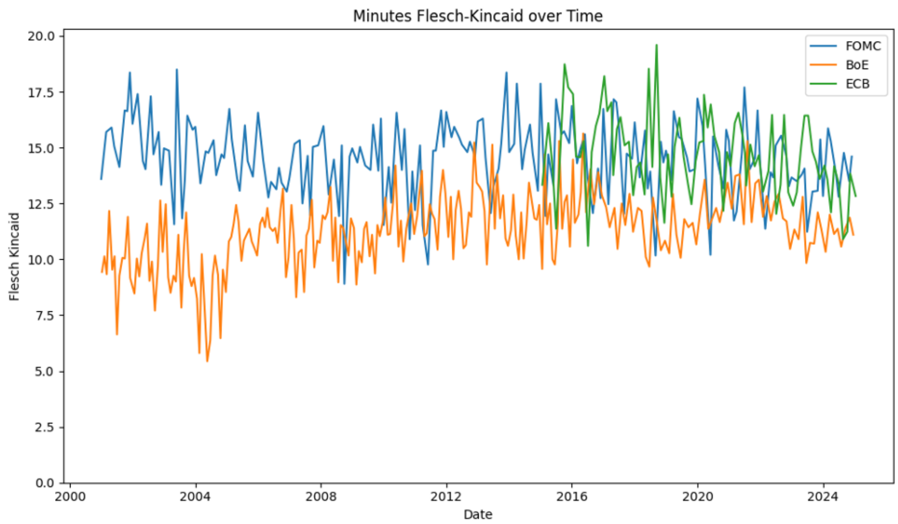

Figure 5: Flesch-Kincaid Score (Minutes)

Based on the Flesch-Kincaid scores for meeting minutes, the following observations are made:

1. BoE minutes appear slightly less complex on average than FOMC and ECB minutes.

2. While statements have experienced some trend-decline in the Flesch-Kincaid score over the past decade, perhaps reflecting efforts to simplify communications, the decline has not materialized to the same extent for minutes.

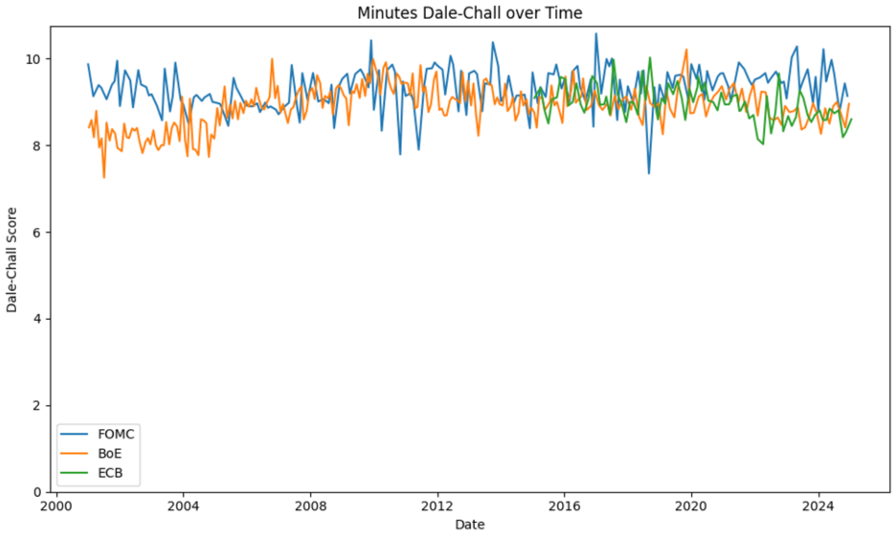

Figure 6: Dale-Chall Score (Minutes)

Based on the Dale-Chall scores for meeting minutes, the following observations are made:

1. The difference in complexity between the three central banks for minutes appears to be far less significant than that of statements.

2. There appears to be a slight increase in complexity following the 2008 Global Financial Crisis for both the FOMC and BoE.

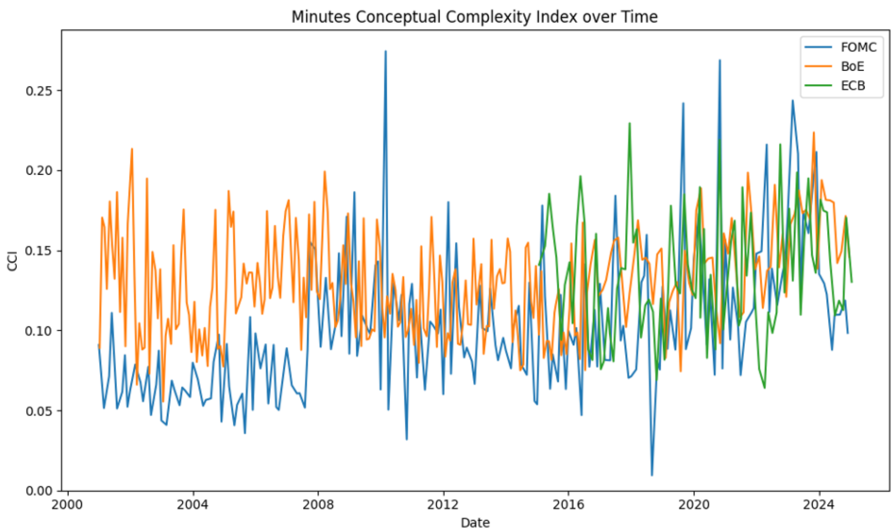

Figure 7: Conceptual Complexity Index (Minutes)

Based on the CCI scores for meeting minutes, the following observations are made:

1. Like the Dale-Chall score, it is difficult to discern any significant and consistent difference in CCI complexity of meeting minutes between the three central banks.

2. The FOMC saw sharp increases in the CCI at certain meetings following the GFC, as well during the post-pandemic inflationary period of the 2020s. The ECB also saw spikes in the CCI in the mid to late 2010s, which coincides with more unconventional actions to support the economy during that time. This also coincides with the rise in CCI scores for ECB statements.

4.3 Measuring Monetary Policy Surprises

In order to understand the relationship between the measured complexity of monetary policy communication and the corresponding change in expected forward-looking volatility in financial markets, the "surprise" component of a monetary policy statement must be included as a control variable. Empirically, monetary policy surprises have a statistically significant impact on financial asset price movements in short time windows.

GSS showed in a high frequency event-study analysis that the movements of asset prices around U.S. monetary policy announcements cannot be adequately captured by a single factor; rather, they require two factors, both of which have a structural interpretation. The first factor can be interpreted as a "target" factor, associated with the surprise component of changes in the target rate. The second factor can be interpreted as a "path" factor, associated with the surprise component of changes in the expected path of future policy. More recently, Swanson (2017) extended the methods of GSS and identified a third factor in the effects of monetary policy post-2009, which can be structurally interpreted as a quantitative easing (QE) factor.

To calculate the factors for use as control variables in our regressions, I follow the approach of Swanson (2017). The starting point is constructing a $T \times n$ matrix of asset price reactions in short-windows around each statement announcement, where there are T announcement days and n assets. The first three principal components that explain a maximal share of the variation in the matrix are then extracted. The three principal components do not have a structural interpretation as target, path, and QE factors. Indeed, it is much more likely that each principal component is some linear comibnation of the three interpretable factors. To address this issue, Swanson (2017) multiplies the matrix of principal components by a 3×3 unitary rotation matrix U, which can be uniquely identified by imposing three constraints. The first two constraints are that the path and QE factor do not load onto the shortest duration asset in the matrix. The third constraint is that the variance of the QE factor is minimal over the period prior to the first instance of QE in each country. The first two constraints are intuitive because surprises regarding the path of monetary policy should have greater effect on longer duration yields than the short-term, and QE involves purchasing long-dated securities, making the effect primarily on the long-end of the yield curve. The third constraint minimizes the QE factor's contribution to the surprises over the period when QE had not yet been implemented. Mathematical details of computing the rotation matrix U can be found in Appendix A.

The following sources provide the data for the matrices of asset price reactions. For the FOMC, the "Monetary Policy Surprises" (MPS) dataset from Bauer and Swanson (2023) was used. For the BoE, the UK Monetary Policy Event-Study Database (UKMPD) from Braun et al (2024) was used. For the ECB, the Euro Area Monetary Policy Event-Study Database (EA-MPD) from Altavilla et al. (2019), which has since been updated to contain data until 2023, was used. Table 1 displays the datasets in greater detail, including the duration of specific securities used.

Table 1: Monetary Policy Surprise Datasets

Central Bank	Dataset Name	Period	Event Types	Securities Used
FOMC	MPS	1988-2023	Scheduled and unscheduled announcements	Current-quarter, next-quarter, and three-quarter ahead Eurodollar futures contracts, 2-year, 5-year, and 10-year Treasury yields
BoE	UKMPD	1977-2024	MPC announcements, press conferences	1st quarterly, 2nd quarterly, and 4th quarterly LIBOR/SONIA contracts, 2-year, 5-year, and 10-year Gilt yields
ECB	EA-MPD	1999-2023	Press Release, Press Conference	3-month, 6-month, 1-year, 2-year, 5-year, and 10-year overnight index swap (OIS) rate

Monetary policy announcements in the ECB differ from the U.S. and U.K. in the sense that the policy decision and the accompanying statement are announced at different times, with the decision typically being only a few sentences long and the statement providing greater detail. The policy decision is typically revealed in a press release 30 minutes before the press conference, during which the statement is given. The EA-MPD includes data on asset price reactions around both the time of policy decision release and the press conference, as well as reactions over a short time window containing both events. Because the regressions detailed in Section 4.4 utilize daily reactions in the volatility instruments, the price reaction over the entire period, which includes both events, is used.

4.4 Assessing Market Reaction to Statement Complexity

Higher complexity, both in semantic and conceptual terms, may potentially hinder the public's understanding of the stance of monetary policy. To test this hypothesis, the relationship between complexity of a central bank statement and forward-looking volatility in financial markets is examined.

A potential hypothesis for how language complexity may affect financial market reactions is the following: if the language used is more complex, market participants may anticipate higher volatility in the near future as disagreements or confusion over the central bank's message are resolved. A way to assess whether this anticipation occurs is to track the price of certain financial derivatives. In most developed-economy markets, traders regularly take positions on instruments that rise if volatility is expected to be higher in the near future, such as the CBOE Volatility Index (VIX) in the U.S. By measuring the price reaction for such instruments around central bank statements, the relationship between language complexity in statements and expected volatility going forward can be evaluated.

The specific instruments analyzed are the VIX for the U.S., the FTSE Implied Volatility Index (IVI) for England, and Euro STOXX Volatility (VSTOXX) for Europe. The VIX is a measure designed to reflect 30-day expected volatility of the U.S. stock market, calculated using real-time prices of call and put options on the S&P 500 Index, which tracks the performance of 500 of the largest companies in the U.S. The IVI is a series of end-of-day indices that reflect the 30-day expected volatility of the FTSE 100, an index that tracks the performance of 100 large companies in the U.K. The VSTOXX reflects market expectations of 30-day volatility in the EURO STOXX 50, which tracks the performance of 50 blue-chip companies from 11 countries in the Eurozone. Like the VIX, both the IVI and VSTOXX are also calculated based on prices of options on the respective indices.For each index, daily price quotes are retrieved from Bloomberg for the entire time duration of the central bank statements sample. Daily intervals were chosen because intraday data is not readily available for the IVI and VSTOXX.

The following regressions are ran for each country in the sample. For each central bank $i \in \{\text{FOMC, BoE, ECB}\}$, the respective volatility measure is regressed on a language complexity score and a vector of factors that summarize the "surprise" contained in the policy announcement. The vector of factors is included to control for the fact that the volatility instruments may react to surprises in the statement, in addition to the way the statement is constructed. The policy surprises—which represent surprises related to change in the target rate, change in the path forward, and change in quantitative easing policies—are not related to the complexity of the statement, but nonetheless must be incorporated into prices. More detail on factor estimation in Appendix A.

Specifically, the regressions are defined as follows

$$v_i = \alpha_i + \beta_i \text{fk}_i + \gamma_i \text{factors}_i$$

$$v_i = \alpha_i + \beta_i \text{dc}_i + \gamma_i \text{factors}_i$$

$$v_i = \alpha_i + \beta_i \text{cci}_i + \gamma_i \text{factors}_i$$

Here, v_i is the percentage change in the level of the volatility index for country i on the statement day, fk_i is the Flesch-Kincaid score for the statement, dc_i is the Dale-Chall score for the statement, and cci_i is the Conceptual Complexity Index for the statement. $factors_i$ is the vector of target, path, and QE factors that represent a decomposition of the monetary policy surprise. α_i is included as a constant.

The effect of a monetary policy surprise on expectations of volatility is not entirely clear. On one hand, an unexpected rate cut could reduce volatility expectations by lowering borrowing costs and fostering more stable economic growth by easing pressures . On the other hand, it might signal underlying economic weakness, leading market participants to revise their outlook and anticipate greater volatility. To account for the possibility that positive and negative monetary policy surprises influence volatility expectations differently—both in magnitude and direction—the same three regressions are ran but with the surprises separated into positive and negative. The regressions are specified as follows:

$$v_i = \alpha_i + \beta_i \mathrm{fk}_i + \gamma_{+,i}\mathrm{factors}_{+,i} + \gamma_{-,i}\mathrm{factors}_{-,i}$$
$$v_i = \alpha_i + \beta_i \mathrm{dc}_i + \gamma_{+,i}\mathrm{factors}_{+,i} + \gamma_{-,i}\mathrm{factors}_{-,i}$$
$$v_i = \alpha_i + \beta_i \mathrm{cci}_i + \gamma_{+,i}\mathrm{factors}_{+,i} + \gamma_{-,i}\mathrm{factors}_{-,i}$$

For each country i and jth term in the vector,the positive and negative components of $\mathrm{factors}_{ij}$ are defined as follows:

$$\mathrm{factors}_{+,ij} = \begin{cases} \mathrm{factors}_{ij}, & \mathrm{factors}_{ij} > 0 \\ 0, & \mathrm{factors}_{ij} \leq 0 \end{cases}$$

$$\mathrm{factors}_{-,ij} = \begin{cases} \mathrm{factors}_{ij}, & \mathrm{factors}_{ij} < 0 \\ 0, & \mathrm{factors}_{ij} \geq 0 \end{cases}$$

5 Results and Discussion

Table 2 displays the regression results for the FOMC. Most of the surprise factors were not significant, suggesting that volatility instruments like the VIX do not necessarily exhibit a linear relationship with surprises in the target rate, path of policy, or quantitative easing. Because the only factor that was significant was the negative path surprise at a 5% level, we drop all other factors from the regressions, leaving only the complexity scores and the negative path surprise as independent variables. (The coefficients remain on the same order of magnitude compared to when the factors are included as controls.[2]) Even after dropping the factors, none of the coefficients are significant. The fact that the Flesch-Kincaid score and Dale-Chall score are insignificant suggest that higher semantic complexity (longer words, longer sentences) do not negatively impact market participants' understanding of FOMC communications. The CCI is also not significant, although it is important to keep in mind that the CCI was developed for use on Bank of England statements, and thus the lexicon may not be appropriate for the

[2] The results for the original regressions, containing the factors as controls, can be found in Appendix D.

FOMC. Finally, the negative path surprise is both significant and positive, indicating that the VIX generally increases in response to such surprises. This may be because negative path surprises reveal previously unknown economic weakness to market participants, leading them to anticipate greater volatility in the future.

Table 3 displays results for the ECB. Consistent with the FOMC results, most of the factors were once again not significant predictors of changes in the volatility instrument (VSTOXX). The only significant factor for all three scores was the positive QE surprise, so all other factors were dropped from the regression specification. The Flesch-Kincaid score is significant at a 1% level, and the CCI is significant at a 5% level. The positive sign of the Flesch-Kincaid suggests that greater semantic complexity leads to an increase in expected volatility as measured by the VSTOXX, suggesting that greater semantic complexity may be confusing to market participants. This is consistent with the results of Bulir (2014), who found evidence that more clarity was associated with lower levels of return volatility in financial markets for the Euro area. On the other hand, the negative sign on the CCI suggests that it may be possible that greater conceptual complexity is helpful for market participants in understanding policy communication, as it lowers the level of the VSTOXX. After dropping all other factors, the positive QE surprise is no longer significant.

To assess whether the ECB regression results have practical relevance beyond statistical significance, I calculate the impact of a one standard deviation change in the complexity scores on VSTOXX. The standard deviation of the ECB's Flesch-Kincaid score is 1.46; when multiplied by the coefficient of 0.013, this implies that a one standard deviation increase in the Flesch-Kincaid score corresponds to VSTOXX rising $\sim$1.9% on the announcement day. The standard deviation of the ECB's CCI is 0.035; when multiplied by the coefficient of -0.290, this implies that a one standard deviation increase in the CCI corresponds to VSTOXX declining $\sim$1.0%. These results suggest that the measured effects are not only statistically significant but also economically meaningful, as they drive notable movements in VSTOXX.

Finally, Table 4 displays results for the BoE. None of the factors were consistently significant across all three scores, so we drop all of them from the regressions. After dropping the factors, none of the complexity scores are significant; this suggests that neither semantic nor conceptual complexity has a negative effect on market participants' understanding of BoE communications,

Table 2: FOMC Regressions

	Central Bank: FOMC		
	FK	DC	CCI
	(1)	(2)	(3)
Intercept	-0.046	-0.034	-0.000
	(0.056)	(0.114)	(0.019)
fk	0.002		
	(0.004)		
dc		0.002	
		(0.012)	
cci			-0.105
			(0.175)
path_	0.382***	0.393***	0.368**
	(0.145)	(0.145)	(0.149)
Observations	119	119	119
R^2	0.063	0.060	0.062
Adjusted R^2	0.046	0.044	0.046
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 3: ECB Regressions

	Central Bank: ECB		
	FK	DC	CCI
	(1)	(2)	(3)
Intercept	-0.172***	-0.152	0.028*
	(0.043)	(0.140)	(0.016)
fk	0.013***		
	(0.003)		
dc		0.017	
		(0.016)	
cci			-0.290**
			(0.123)
QE_+	-0.005	-0.004	-0.003
	(0.004)	(0.004)	(0.004)
Observations	200	200	200
R^2	0.075	0.012	0.034
Adjusted R^2	0.066	0.002	0.024
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 4: BoE Regressions

	Central Bank: BoE		
	FK	DC	CCI
	(1)	(2)	(3)
Intercept	0.144	-0.331	-0.055
	(0.156)	(0.454)	(0.054)
fk	-0.011		
	(0.012)		
dc		0.039	
		(0.052)	
cci			0.310
			(0.263)
Observations	38	38	38
R^2	0.021	0.015	0.037
Adjusted R^2	-0.006	-0.012	0.010
Note:	*p<0.1; **p<0.05; ***p<0.01		

but also could be a result of the BoE sample being smaller.

6 Concluding Remarks

This thesis analyzed the language complexity of statements and minutes for three central banks: the FOMC, the ECB, and the BoE. For each document, complexity was measured using two *semantic* scores (Flesch-Kincaid and Dale-Chall) and one *conceptual* score (Conceptual Complexity Index). While FOMC statements appear to be the most complex from a semantic standpoint, the BoE appears to be most complex from a conceptual standpoint.

This thesis also investigated the effect of language complexity—as measured by the three scores—on changes in forward-looking volatility, in an effort to understand how language complexity might affect market participants. The effects are found to vary across the three central banks, with language complexity having the most pronounced impact in the ECB, where both the Flesch-Kincaid score and CCI are significant predictors of the volatility instrument. Notably, while higher Flesch-Kincaid scores lead to an increase in forward-looking volatility, higher CCI actually leads to a decrease in forward-looking volatility. Specifically, a one standard deviation increase in the Flesch-Kincaid score causes the European implied volatility index (VSTOXX) to rise $\sim$1.9% on the announcement day, and a one standard deviation increase in

CCI causes VSTOXX to decrease by $\sim$1.0%.[3]

These results suggest that higher conceptual complexity, characterized by increased use of economic jargon and larger breadth of topics, may be helpful for market participants' understanding of monetary policy. Nevertheless, the impact of complex language on market participants is statistically insignificant for both the FOMC and the BoE. At the same time, however, being understood by financial markets may not directly translate into being understood by the general public. Future research could further investigate the effects of complex language on the public at large. Future research could also explore language complexity in other types of monetary policy documents. Speeches and Q&A sessions, in particular, are interesting because the public receives them aurally rather than through reading.

References

Altavilla, C., Brugnolini, L., Gürkaynak, R. S., Motto, R., & Ragusa, G. (2019). Measuring euro area monetary policy. *Journal of Monetary Economics*, 108, 162–179. https://doi.org/10.1016/j.jmoneco.2019.04.004

Amato, J. D., Morris, S. E., & Shin, H. S. (2003). Communication and monetary policy. *BIS Working Paper No. 123*. https://ssrn.com/abstract=845444

Bank of England. (n.d.). *Bank of England*. https://www.bankofengland.co.uk

Bauer, M. D., & Swanson, E. T. (2023). Monetary policy surprises dataset. In *A reassessment of monetary policy surprises and high-frequency identification. NBER Macroeconomic Annual*, 23. National Bureau of Economic Research. https://www.frbsf.org/research-and-insights/data-and-indicators/monetary-policy-surprises/

Blinder, A. S., 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. https://doi.org/10.1257/jel.46.4.910

Bloomberg L.P. (2024). VIX, IVI, and VSTOXX daily price quotes. Retrieved from Bloomberg database.

Braun, R., Miranda-Agrippino, S., & Saha, T. (2023). Measuring monetary policy in the UK:

[3] The code used for this thesis can be made available upon request at alexander.ye21@gmail.com.

The UK Monetary Policy Event-Study Database. *Staff Working Papers*. Bank of England.

Bulíř, A., Čihák, M., & Jansen, D.-J. (2014). Does the clarity of inflation reports affect volatility in financial markets? *IMF Working Paper*, Institute for Capacity Development.

Chall, J. S., & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Brookline Books.

Dale, E., & Chall, J. S. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27(1), 11–20+28. https://www.jstor.org/stable/1473169

Danker, D. J., Luecke, M. M. (2005). Background on FOMC meeting minutes. Federal Reserve Bulletin, 91(Spring), 175–179.

de Haan, J., & Hoogduin, L. (2024). ECB communication policies: An overview and comparison with the Federal Reserve. *Journal of International Money and Finance*, 142, C. Elsevier.

European Central Bank. (n.d.). *European Central Bank*. https://www.ecb.europa.eu

Federal Reserve. (n.d.). *Federal Reserve*. https://www.federalreserve.gov

Federal Reserve History. (2024, August 5). Transparency. *Federal Reserve History*. Jonathan Rose contributed to this article.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233. https://doi.org/10.1037/h0057532

Gardt, M., Angino, S., Mee, S., & Glöckler, G. (2021). ECB communication with the wider public. *ECB Economic Bulletin*, Issue 8/2021.

Greenspan, A. (1993). Testimony before the Committee on Banking, Finance, and Urban Affairs, U.S. House of Representatives, October 13, 1993.s

Gürkaynak, R., Sack, B., & Swanson, E. (2005). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, 1(1), 55–93.

Jefferson, P. N. (2024, May 13). Communicating about monetary policy. Speech presented at the "Central Bank Communications: Theory and Practice" conference, Federal Reserve Bank of Cleveland, Cleveland, Ohio.

Kincaid, J. P., Fishburn, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas for Navy enlisted personnel (pp. 8-75). *Technical Report Research*

Branch Report. Millington, TN: Naval Technical Training. Memphis, TN: US Naval Air Station.

Kuttner, K. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics*, 47(3), 523–544.

Macklem, T., & Vardy, J. (2023). 20 years of central bank communications, and lessons for the future. In *Central banking in the Americas: Lessons from two decades* (Vol. 127, pp. 55–67). Bank for International Settlements.

McMahon, M., & Naylor, M. (2023). Getting through: Communicating complex information. *Staff Working Papers*, Bank of England.

Millard, S. (2022). Central bank communication: Never excuse, never explain (*NIESR Policy Paper 33*). National Institute of Economic and Social Research.

Nakamura, E., & Steinsson, J. (2013, revised 2018). High-frequency identification of monetary non-neutrality: The information effect (*Working Paper No. 19260*). National Bureau of Economic Research. https://www.nber.org/papers/w19260

Ragusa, G., Motto, R., Brugnolini, L., Gürkaynak, R., & Altavilla, C. (2019, October 3). The Euro Area Monetary Policy Event-Study Database. *CEPR VoxEU*. https://cepr.org/voxeu/columns/euro-area-monetary-policy-event-study-database

Smales, L. A., & Apergis, N. (2016). Does more complex language in FOMC decisions impact financial markets? *Asian Finance Association (AsianFA) 2016 Conference, Journal of International Financial Markets, Institutions and Money*, 51, 1–14. https://doi.org/10.2139/ssrn.2724364

Swanson, E. T. (2017). Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets (*Working Paper No. 23311*). National Bureau of Economic Research. https://doi.org/10.3386/w23311

Swanson, E. T., & Jayawickrema, V. (2024). Speeches by the Fed Chair are more important than FOMC announcements: An improved high-frequency measure of U.S. monetary policy shocks (*Working Paper*).

Woodford, M. (2005). Central bank communication and policy effectiveness (*NBER Working Paper No. 11898*). National Bureau of Economic Research. https://doi.org/10.3386/w11898

Yellen, J. L. (2012, November 13). Revolution and evolution in central bank communications. Speech at the Haas School of Business, University of California, Berkeley, Berkeley, California.

7 Appendix A: Estimating Factors

This section explains the process of identifying factors from the data, following the methodology introduced by Swanson (2017), which identifies three factors that can be interpreted as a 'target' surprise, a 'path' surprise, and a 'QE' surprise.

For each central bank, the starting point is constructing a $T \times n$ matrix of yield surprises in short windows around announcements. To maintain consistency across all three central banks, I select a set of $n = 6$ yield responses from the monetary policy surprise datasets that are equal in duration. For the FOMC, the selected asset price responses were the current-quarter Eurodollar (ED) futures contract, the next-quarter ED futures contract, and the three-quarter-ahead ED futures contract, as well as the 2-year, 5-year, and 10-year Treasury yield responses. For the Bank of England, the selected responses were the 1st, 2nd, and 4th quarterly contracts of the 3-month LIBOR, as well as the 2-year, 5-year, and 10-year Gilt yields. (Since 2021, the SONIA rate is used instead of LIBOR.) For the ECB, the selected responses were the 3-month, 6-month, 1-year, 2-year, 5-year, and 10-year overnight index swap (OIS) rates. Because data for OIS rates at maturities longer than 2-years is only available after August 2011, prior to that date German bond yields are used as a proxy, following Altavilla et al. (2019).

Ultimately, the aforementioned securities were chosen because they roughly correspond to yields with duration of 3-month, 6-month, 1-year, 2-year, 5-year, and 10-year, respectively. The columns of the $T \times n$ matrix are arranged such that the duration increases going from the leftmost column to the right. In total, the FOMC matrix contains $T = 296$ announcements, the BoE matrix contains $T = 288$ announcements, and the ECB matrix contains $T = 277$ announcements.

After constructing the the $T \times n$ matrix X, each column of X is normalized to have zero mean and unit variance. I then estimate the three factors that explain a maximal fraction of the variance of the data using principal component analysis, according to the following model:

$$X = F\Lambda + \varepsilon$$

F is a $T \times 3$ matrix representing the three factors, where the columns are orthogonal. Λ is a $3 \times n$ matrix representing the factor loadings. ε is an error term, which is minimized. The columns of F cannot be interpreted as monetary policy surprises unless certain restrictions are imposed on them; indeed, there is no reason why the first column of F should be correlated with the target, path, or QE surprise. (In reality, each column of F is more likely to be some linear combination of the three factors.) The goal is to estimate $F^* = FU$, where U is a 3×3 orthogonal matrix that rotates F in such a way that the columns of F^* can be interpretable as a 'target', 'path', and 'QE' factor. Because the only transformation applied to F is a rotation, the new matrix F^* will explain the same amount of variance in X as F.

Because U is a 3×3 orthogonal matrix, it requires three identifying restrictions. The three restrictions outlined by Swanson (2017) are the following:

1. The path factor has no effect on the current-month federal funds rate.

2. The QE factor has no effect on the current-month federal funds rate.

3. The variance of the QE factor is minimized over the pre-ZLB period.

The first and second assumptions are intuitive because surprises concerning the path of monetary policy or QE should primarily affect medium- and long- term yields but not necessarily short-term yields. This is because path is defined as the future trajectory of monetary policy, while QE is conducted by purchasing securities on the long end of the yield curve. The third assumption distinguishes between the path and QE factor, with the idea being that prior to the Great Financial Crisis, QE should have minimal impact on yield changes.

Since the shortest duration security in my $T \times n$ matrix is the 3-month yield, I modify Swanson's restriction and require that the path and QE factor have no effect on the 3-month yield. Furthermore, the pre-ZLB period is defined as before November 2008 for the FOMC, before March 2009 for the BoE, and before January 2015 for the ECB. These dates mark the first instances of QE for each of the central banks, respectively.

Swanson's first two restrictions can be mathematically defined as

$$U'\Lambda_1 = \begin{pmatrix} \cdot \\ 0 \\ 0 \end{pmatrix}$$

where Λ_1 is the first column of Λ, corresponding to the factor loadings onto the shortest duration yield (first column of X). This ensures that only the first factor will be correlated with the shortest duration yield. The third restriction is that the variance of the QE factor is minimized prior to the pre-ZLB date, as defined earlier. Since the QE factor is given by FU_3, this is equivalent to minimizing $U_3'(F^{pre})'(F^{pre})U_3$, where F^{pre} includes all the announcement dates of F that are within the pre-ZLB period.

Swanson showed that implementing these three restrictions is equivalent to solving the following objectives and constraints:

1. Minimize
$$\begin{bmatrix} u_{13} & u_{23} & 1 \end{bmatrix} (F^{pre})'F^{pre} \begin{bmatrix} u_{13} \\ u_{23} \\ 1 \end{bmatrix}$$
subject to $\Lambda_1' \begin{bmatrix} u_{13} & u_{23} & 1 \end{bmatrix}' = 0$. Rescale the minimizing vector $\begin{bmatrix} u_{13} & u_{23} & 1 \end{bmatrix}'$ to have length 1, and call the resulting vector U_3.
2. Solve
$$\begin{pmatrix} \Lambda_1' \\ U_3' \end{pmatrix} \begin{pmatrix} u_{12} \\ u_{22} \\ 1 \end{pmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
and rescale the solution vector $\begin{bmatrix} u_{12} & u_{22} & 1 \end{bmatrix}'$ to have length 1. Call the resulting vector U_2.
3. Solve
$$\begin{bmatrix} U_2' \\ U_3' \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
and rescale the solution vector $\begin{bmatrix} u_{11} & u_{12} & 1 \end{bmatrix}'$ to have length 1. Call the resulting vector U_1.

The first equation ensures that the QE factor has minimum variance over the pre-ZLB period, subject to the constraint that the QE factor does not load onto the shortest duration yield. The second equation ensures that the path factor does not load onto the shortest duration yield, and also that the path factor is orthogonal to the QE factor. The third equation ensures that the target factor is orthogonal to both the path and QE factor. Having identified the rotation matrix U, I compute $F^* = FU$, where the first, second, and third columns correspond to the target, path, and QE factors.

Finally, the target factor is rescaled to have a positive one-for-one effect on the 3-month yield, the path factor to have a positive one-for-one effect on the 1-year yield, and the QE factor to have a negative one-for-one effect on the 10-year yield.

To visualize the factors, the asset price surprises were regressed on the target, path, and QE factor. Plots of the regression coefficients for each factor, as well as the 90% confidence intervals, are displayed in Figure 8-10.

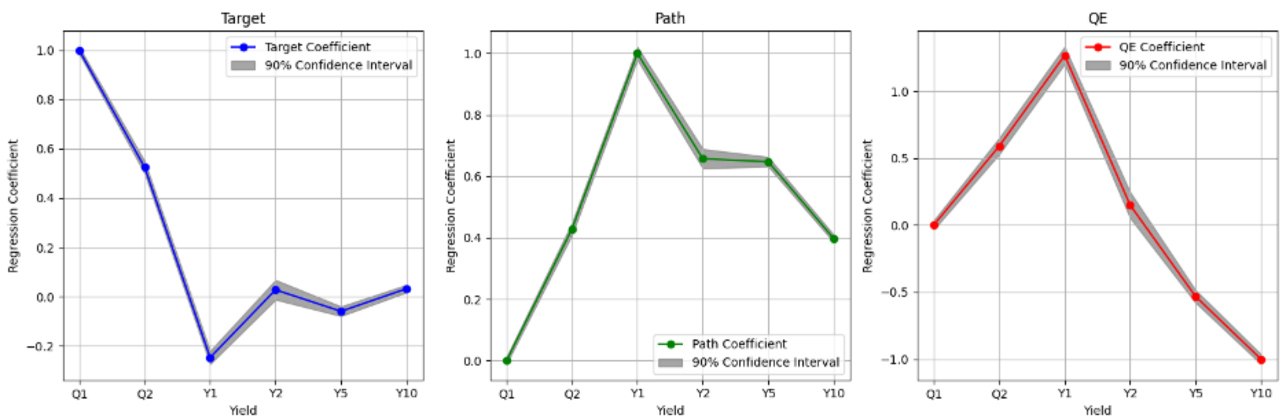

Figure 8: FOMC Factors

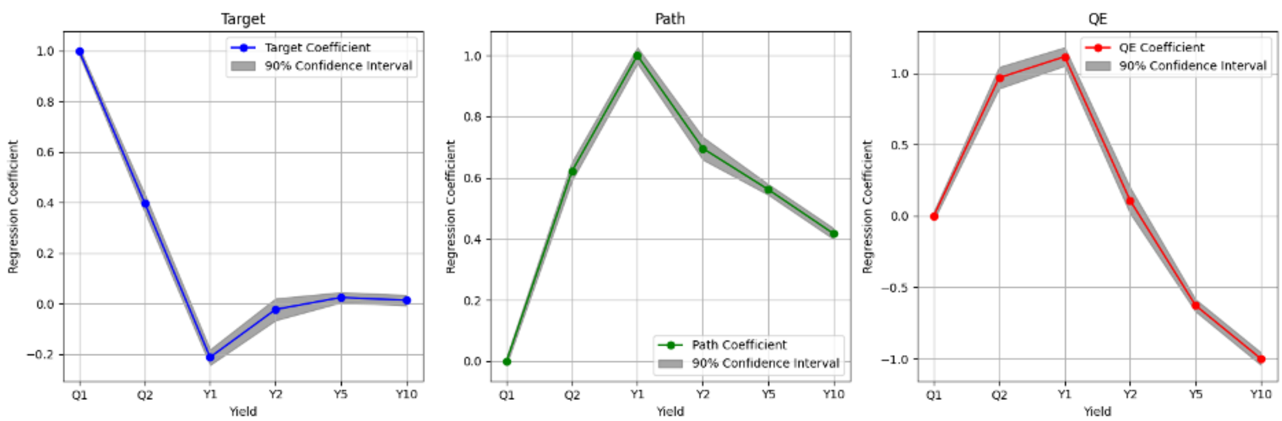

Figure 9: BoE Factors

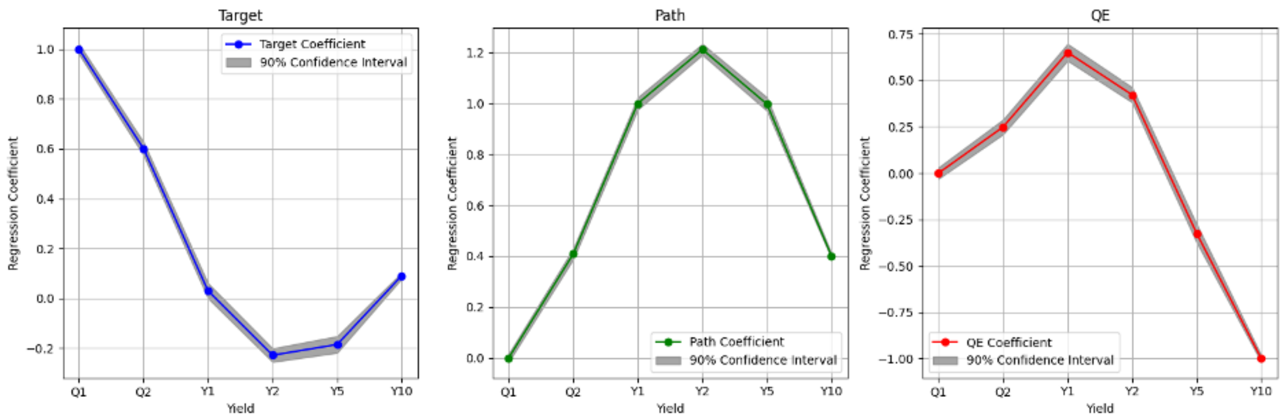

Figure 10: ECB Factors

The target factor is most prominent in the shortest duration yield, and the influence of the target factor decreases as the duration increases. A positive surprise in the target rate increases short-term yields, all else equal. The path factor is most prominent for durations in the medium term of around one to two years. A positive surprise in the path factor increases medium-term yields, all else equal. The QE factor is most prominent for long-term yields. A positive surprise in the QE factor will decrease long-term yields, all else equal. Consistent with Swanson (2017), an increase in QE also causes short-yields to rise slightly, on average.

During the pandemic, central banks purchased securities across the curve to improve market functioning, potentially amplifying the QE factor's effects on securities of all durations. To address this issue, when regressing the language complexity scores on the factors, the regression sample is limited to end in 2019, using only pre-pandemic estimated factors.

8 Appendix B: Sample Texts

This section provides examples of "complex" excerpts from central bank statements, as measured by the three complexity scores.

8.1 Flesch-Kincaid Score

The June 2014 FOMC statement has the highest Flesch-Kincaid score in the entire sample. Below are two comparable excerpts, highlighting factors that cause fluctuations in the Flesch-Kincaid score: the first from the June 2014 statement and the second from the preceding FOMC statement in April 2014.

Excerpt from June 2014 FOMC Statement

"Labor market indicators generally showed further improvement. The unemployment rate, though lower, remains elevated. Household spending appears to be rising moderately and business fixed investment resumed its advance, while the recovery in the housing sector remained slow. Fiscal policy is restraining economic growth, although the extent of restraint is diminishing."

Excerpt from April 2014 FOMC Statement

"Labor market indicators were mixed but on balance showed further improvement. The unemployment rate, however, remains elevated. Household spending appears to be rising more quickly. Business fixed investment edged down, while the recovery in the housing sector remained slow. Fiscal policy is restraining economic growth, although the extent of restraint is diminishing."

Although the excerpts share similar wording and content, the June 2014 excerpt has a Flesch-Kincaid score nearly 2 units higher than the April 2014 excerpt. This increase is mainly due to the June 2014 statement combining the sentence about household spending with the one about business fixed investment and the housing sector. The resulting longer sentence, along with the substitution of certain words with ones containing more syllables, causes the higher Flesch-Kincaid score.

8.2 Dale-Chall Score

The August 2002 FOMC statement has the highest Dale-Chall score in the entire sample, with a score nearly 2 units higher than the preceding June 2002 statement. Below are two comparable sentences that contributed significantly to this difference.

Excerpt from August 2002 FOMC Statement

"The softening in the growth of aggregate demand that emerged this spring has been prolonged in large measure by weakness in financial markets and heightened uncertainty related to problems in corporate reporting and governance."

Excerpt from June 2002 FOMC Statement

"However, both the upward impetus from the swing in inventory investment and the growth in final demand appear to have moderated."

While both sentences describe the slowdown in demand growth, the sentence from August 2002 is significantly longer and includes terms such as “softening,” “aggregate,” “prolonged,” “corporate,” “reporting,” and “governance,” which are not part of the Dale-Chall dictionary of commonly used English words.

One notable finding is that the June 2014 statement, which has the highest Flesch-Kincaid score in the sample, only has a Dale-Chall score of 8.9, which is not particularly high when compared to the other statements. A key sentence that exemplifies why the difference occurs is the following:

“If incoming information broadly supports the Committee’s expectation of ongoing improvement in labor market conditions and inflation moving back toward its longer-run objective, the Committee will likely reduce the pace of asset purchases in further measured steps at future meetings.”

Words like “information”, “expectation”, “improvement”, “objective”, and “committee” cause a higher Flesch-Kincaid score because they have high syllable counts. However, these words are not necessarily rare or difficult words, and many of them fall within the Dale-Chall dictionary of most common words. This example highlights the distinction between the Flesch-Kincaid score and the Dale-Chall score.

8.3 Conceptual Complexity Index

The March 2024 BoE statement has the highest CCI score in the entire sample. Below is an excerpt from the statement that displays key characteristics that cause the CCI to be high.

Excerpt from March 2024 BoE Statement

“Since the MPC’s previous meeting, market-implied paths for advanced economy policy rates have shifted up. In the United States and the euro area, inflationary pressures have continued to abate, though by slightly less than expected. Material risks remain, notably from developments in the Middle East including disruption to shipping through the Red Sea. Having declined through the second half of last year, UK GDP and market sector output are expected to start growing again during the first half of this year.”

Recall that the CCI penalizes both the use of many different jargon words and the coverage of many different economic topics. This paragraph does both: it uses many different jargon words such as “inflationary”, “GDP”, and “output”, and it also covers a wide range of topics including inflation, output, open economy, and financial markets.

9 Appendix C: Conceptual Complexity Index

This section walks through the method for calculating the Conceptual Complexity Index (CCI) developed by McMahon and Naylor (2023). As mentioned, the CCI is designed to be increasing in the proportion of jargon words in the text, the number of distinct jargon terms used within a given topic, and the number of different topics covered. To achieve this, McMahon and Naylor (2023) consider that discussion of the economy involves jargon terms j that can be broadly grouped into T topics.[4] Each jargon term j is mapped to a topic $t \in \{1, \cdots, T\}$. For each document d, the authors first find $W_{j,t,d}$, which is the total number of jargon terms j under topic t.

In designing the CCI, the authors aim to penalize not only texts with a high raw jargon count but also those that use many distinct jargon terms. For instance, if two documents each have 10 jargon terms, but one repeats the same term while the other uses 10 distinct terms, the latter should have a higher CCI. To capture this dynamic, the authors compute a weight $\psi_{t,d}$ which adjusts the jargon count within topic t of document d depending on the “dispersion” of distinct jargon terms used within a topic, calculated as follows:

$$\psi_{t,d} = \sqrt{\sum_{j_t=1}^{J_t} s_{j,t,d}^2}$$

Here, $s_{j,t,d} = \frac{w_{j,t,d}}{W_{j,t,d}}$, where $w_{j,t,d}$ is the number of instances of jargon term $j_t \in \{1, \cdots, J_t\}$ in topic t for document d. (Therefore, we have $W_{j,t,d} = \sum_{j_t=1}^{J_t} w_{j,t,d}$, as $W_{j,t,d}$ is the total number of jargon terms under topic t mentioned in the document.) Essentially, $s_{j,t,d}$ is the proportion of references to a certain jargon term within a topic, out of all jargon references to that topic. For example, if the topic were “economic growth” and the only jargon term used was “GDP”, then $s_{j,t,d} = 1$ for $j_t =$ GDP and $s_{j,t,d} = 0$ for all other terms within the “economic growth” topic. Hence, $\psi_{t,d} = 1$ if only one unique jargon term within topic t is used, and decreases

[4] The $T = 10$ topics identified by McMahon and Naylor (2023) are monetary policy, inflation, output, private demand, fiscal policy, open economy, labor markets, financial markets, financial stability, and other. Other is a broad category that captures all jargon words not contained within the other topics.

towards 0 as the number of unique jargon terms within the topic increases.

The weight $\psi_{t,d}$ is then used to adjust $W_{j,t,d}$, the total number of jargon terms in the topic. The adjusted jargon count is computed as follows: $W^*_{j,t,d} = \frac{W_{j,t,d}}{\Psi_{t,d}}$ where $\Psi_{t,d} = 2^{\log_{10}\psi_{t,d}}$. A transformed version of $\psi_{t,d}$ is used so that if only one unique jargon term under topic t is used, the adjusted jargon count $W^*_{j,t,d}$ is equal to the raw jargon count $W_{j,t,d}$. A decrease of $\psi_{t,d}$ by a factor of 10 will double the adjusted jargon count. To summarize, $W^*_{j,t,d}$ is increasing both in the total amount of jargon terms under topic t and in the total number of different jargon terms under topic t.

The CCI also aims to penalize documents that cover many different topics. McMahon and Naylor (2023) design the following weight Φ_d, which follows from the term frequency - inverse document frequency (tf-idf) weighting commonly used in natural language processing.

$$\Phi_d = \frac{\log_{10}(T_d + v)}{\log_{10}(T_d + v) - log_{10}T_d}$$

Here, T_d is the number of different topics covered in document d and v is a coefficient that adjusts the level of penalization for additional topics. The weight is increasing in T_d. The authors select $v = 90$ so that if all topics are covered, the weight Φ_d is doubled, compared to a baseline of one topic. [5]

Putting it all together, the CCI is calculated according to the following formula:

$$\text{CCI}_d = \frac{\left(\sum_{i=1}^{T} W^*_{j,t,d}\right) \times \Phi_d}{W_{i,d}}$$

The adjusted jargon counts $W^*_{j,t,d}$ are summed over all possible topics, and then multiplied by the topic-coverage weight Φ_d. The result is then divided by the total number of words $W_{i,d}$ in document d. The score thus exhibits all three features that the authors intend, increasing in total jargon words used, number of different jargon words used, and number of different topics covered.

To quantify the magnitude of changes needed to increase the CCI by approximately 5%, first consider the case where 8 different topics are covered; this results in $\Phi_d \approx 1.83$. Adding

[5] When only 1 topic is covered, $\Phi_d = \frac{\log_{10}(1+90)}{\log_{10}(1+90)-\log_{10}1} = 1$. If all 10 total topics are covered, then $\Phi_d = \frac{\log_{10}(10+90)}{\log_{10}(10+90)-\log_{10}10} = 2$.

an additional topic yields $\Phi_d \approx 1.92$, an increase in Φ_d by 4.9%. Therefore, holding all else equal, a document with 9 different topics will have a CCI approximately 4.9% higher than a document with 8 different topics.

Another way to increase the CCI by 5% is to change the total number of jargon terms. Increasing $W_{j,t,d}$, the number of jargon terms under topic t, by 5% for all T topics will increase $\sum_{i=1}^{T} W^*_{j,t,d}$ by 5%, holding the disperson of jargon terms equal. Hence, the CCI will be 5% higher if the total number of jargon terms increases by 5%.

Reducing the transformed "disperson" weight $\Psi_{t,d}$ by approximately 5% for all topics will also increase the CCI by 5%. The amount of change in specific words needed depends on the number of distinct jargon terms in the document, but adding more unique terms will lower the dispersion weight.

10 Appendix D: Regression Results with Factors

This section displays the original regression results, with factors included as controls.

Table 5: FOMC Regressions with Factors

	Central Bank: FOMC					
	FK	DC	CCI	FK (Separated Surprises)	DC (Separated Surprises)	CCI (Separated Surprises)
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-0.060	-0.032	-0.011	-0.049	-0.020	-0.004
	(0.056)	(0.114)	(0.020)	(0.060)	(0.118)	(0.021)
fk	0.002			0.002		
	(0.004)			(0.004)		
dc		0.001			0.001	
		(0.012)			(0.013)	
cci			-0.110			-0.099
			(0.174)			(0.186)
target	0.137	0.136	0.123			
	(0.228)	(0.229)	(0.229)			
path	0.243	0.250	0.241			
	(0.166)	(0.167)	(0.167)			
QE	0.423	0.443	0.431			
	(0.418)	(0.418)	(0.417)			
target_{-}				-0.164	-0.130	-0.166
				(0.368)	(0.367)	(0.369)
target_{+}				0.227	0.197	0.221
				(0.309)	(0.314)	(0.309)
path_{-}				0.580**	0.585**	0.582**
				(0.278)	(0.278)	(0.278)
path_{+}				0.016	0.004	-0.004
				(0.248)	(0.248)	(0.247)
QE_{-}				0.614	0.555	0.468
				(0.851)	(0.850)	(0.859)
QE_{+}				0.616	0.649	0.669
				(0.497)	(0.495)	(0.495)
Observations	119	119	119	119	119	119
R^2	0.074	0.070	0.074	0.094	0.091	0.093
Adjusted R^2	0.041	0.038	0.041	0.037	0.034	0.036

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6: ECB Regressions with Factors

	Central Bank: ECB					
	FK	DC	CCI	FK (Separated Surprises)	DC (Separated Surprises)	CCI (Separated Surprises)
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-0.176***	-0.165	0.028*	-0.161***	-0.181	0.021
	(0.043)	(0.140)	(0.016)	(0.044)	(0.140)	(0.018)
fk	0.013***			0.012***		
	(0.003)			(0.004)		
dc		0.017			0.020	
		(0.016)			(0.016)	
cci			-0.308**			-0.245*
			(0.122)			(0.128)
target	0.004	0.004	0.004			
	(0.003)	(0.003)	(0.003)			
path	-0.003	-0.003	-0.004*			
	(0.002)	(0.002)	(0.002)			
QE	-0.005*	-0.005	-0.005			
	(0.003)	(0.003)	(0.003)			
target_-				0.002	0.001	0.002
				(0.003)	(0.003)	(0.003)
target_+				0.006	0.007*	0.007
				(0.004)	(0.004)	(0.004)
path_-				-0.004	-0.004	-0.004
				(0.002)	(0.003)	(0.003)
path_+				-0.003	-0.003	-0.003
				(0.003)	(0.003)	(0.003)
QE_-				-0.001	0.000	-0.001
				(0.004)	(0.004)	(0.004)
QE_+				-0.009**	-0.009**	-0.008*
				(0.005)	(0.005)	(0.005)
Observations	200	200	200	200	200	200
R^2	0.084	0.021	0.047	0.099	0.050	0.061
Adjusted R^2	0.065	0.001	0.027	0.066	0.016	0.027

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: BoE Regressions with Factors

	Central Bank: BoE					
	FK	DC	CCI	FK (Separated Surprises)	DC (Separated Surprises)	CCI (Separated Surprises)
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.200	-0.409	-0.062	0.277	-0.385	-0.086
	(0.159)	(0.446)	(0.053)	(0.185)	(0.462)	(0.056)
fk	-0.015			-0.023		
	(0.013)			(0.015)		
dc		0.048			0.043	
		(0.051)			(0.053)	
cci			0.366			0.443
			(0.259)			(0.282)
target	-0.743	-0.598	-0.292			
	(1.428)	(1.431)	(1.416)			
path	1.028	0.781	0.722			
	(0.877)	(0.847)	(0.831)			
QE	0.648	0.177	0.153			
	(1.594)	(1.542)	(1.516)			
target_-				1.412	-0.060	1.741
				(2.525)	(2.511)	(2.584)
target_+				-2.339	-0.908	-1.690
				(2.206)	(2.102)	(2.076)
path_-				-0.497	-0.434	-0.734
				(1.266)	(1.311)	(1.266)
path_+				2.682*	2.132	2.473*
				(1.462)	(1.464)	(1.434)
QE_-				2.494	2.511	3.179
				(3.150)	(3.237)	(3.163)
QE_+				-0.622	-1.377	-1.469
				(2.002)	(1.974)	(1.913)
Observations	38	38	38	38	38	38
R^2	0.154	0.140	0.167	0.230	0.188	0.233
Adjusted R^2	0.052	0.036	0.067	0.050	-0.001	0.054

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$