Complexity in FOMC Communication: Drivers, Differences, and Market Reactions

Tran Hoang Phuong Linh*

April 29, 2026

Abstract

This empirical paper studies the determinants and consequences of linguistic complexity in Federal Open Market Committee (FOMC) communication using a large corpus of policymaker speeches. First, do policymakers adjust their language in response to financial-market conditions before they speak? Complexity responds primarily to volatility-based indicators and trading volume—rather than to returns themselves. Policymakers simplify their language when equity-market volatility rises, but increase structural and syntactic complexity when uncertainty originates in Treasuries, gold, or volatility-linked assets. Second, do individual policymakers differ systematically in how they communicate? Speaker characteristics and topic choice explain most of the variation in complexity, with education, professional background, institutional role, and intellectual tradition all leaving distinct linguistic signatures. Third, does complexity influence market behaviour after the speech? Complexity predicts volatility and trading volume across equities, Treasuries, gold, and volatility indices, with short-lived intraday spikes but overall dampening effects at the daily horizon. Together, the results show a two-way interaction: market turbulence shapes the complexity of FOMC communication, and complexity in turn moderates market responses once information is absorbed.

JEL classification:
Key Words:

*McGill University, Canada. E-mail: linh.tran2@mail.mcgill.ca

1 Introduction

Central bank communication is increasingly recognized as a core policy instrument, not merely a supplement to interest rate decisions but a vital channel through which monetary policy makers shape expectations and beliefs. As communication has become more central to the conduct of monetary policy, research on this topic has expanded rapidly. This growth has been facilitated by advances in natural language processing tools and by the systematic digitization of speeches, statements, and other textual outputs produced by central banks.

A large share of this literature examines how central bank communication affects financial markets and macroeconomic outcomes, with a particularly strong emphasis on the former. Numerous empirical studies demonstrate that communication plays a crucial role in enhancing the predictability of monetary policy decisions and in supporting the broader objectives of central banks (Blinder et al., 2008). Earlier work in this area relied on relatively small corpora and manual coding, reflecting the difficulty of working with textual data. More recent contributions, enabled by modern computational tools, have shifted toward finer-grained analyses of tone, sentiment, policy preferences (such as hawkishness or dovishness), and forward guidance. Measures of textual complexity—such as volume-based measures and readability—have started to gain attention in recent years, albeit much less studied compared to other aspects, as central banks increasingly emphasize accessibility and public understanding. Initiatives such as the Federal Reserve's Fed Listens program and the European Central Bank's explicit commitment to reducing complexity underscore the growing importance of this dimension.

Despite this progress, less is known about the determinants of communication itself. Most existing studies implicitly treat communication as homogeneous or institutionally determined, even though it is produced by individuals with distinct backgrounds, incentives, and personal communication styles. Moreover, communication may respond to prevailing economic conditions, especially during periods of heightened stress when policymakers face greater uncertainty and scrutiny. While some research has explored how tone or sentiment varies with economic conditions, systematic evidence on how textual complexity responds to stress remains limited. Even less is known about whether and how individual characteristics, such as educational or professional background, shape the way policymakers communicate in terms of complexity.

Understanding individual heterogeneity in communication is important for several reasons. First, FOMC members differ widely in training, career experience, and professional norms, which may shape how they process information and articulate policy views. Second, markets are increasingly attentive to individual speeches for signals about policy direction, making variation in communication style potentially consequential for expectations and uncertainty. Finally, if communication complexity systematically reflects personal background rather than institutional strategy, this has implications for transparency, accountability, and the design of communication guidelines within the Federal Reserve System.

This paper addresses key gaps in our understanding of central bank communication by analyzing the speeches of Federal Open Market Committee (FOMC) members and examining three central questions. First, do policymakers adjust the complexity of their communication in response to financial-market conditions observed prior to the speech? The results show that complexity responds most strongly to volatility-based indicators—such as high–low price ranges, absolute returns, and attention measures like trading volume—rather than to simple or abnormal returns. Members simplify their language when broad equity-market volatility rises, but increase structural and syntactic complexity when uncertainty originates in other asset classes such as Treasuries, gold, or volatility-linked instruments. Second, do individual policymakers differ systematically in their communication styles, and are these differences linked to observable characteristics such as education, professional background, institutional role, or topic choice? The analysis reveals substantial and systematic heterogeneity: speaker backgrounds and speech topics account for the majority of variation in complexity, with gender, race, academic training, pre-Fed career, freshwater–saltwater affiliation, and institutional roles all shaping communication patterns in meaningful ways. Third, does the complexity of FOMC communication influence financial-market behaviour in the period following the speech? The findings indicate that speech complexity has statistically significant predictive power for market volatility and trading volume across both equities and Treasuries. These effects can emerge within minutes in intraday data and are strongest for volume-based measures, suggesting that linguistic complexity not only responds to market conditions but also shapes how markets process and react to policy communication.

2 Literature Review

The literature on central bank communication has expanded rapidly over the past two decades, particularly empirical work examining how central bank statements and speeches affect financial markets and the broader economy. The first subsection reviews this broader literature and situates the paper's research questions within it, highlighting that linguistic complexity remains a relatively new but increasingly important dimension of communication, and that the determinants of speech content—how communication is shaped by economic or institutional conditions—are still comparatively underexplored. The second subsection then turns to studies focusing specifically on FOMC members' speeches, summarizing what is known about communication heterogeneity across individuals while underscoring that existing work has largely emphasized policy preferences or institutional roles rather than linguistic complexity. Together, these sections clarify how the paper contributes to both strands of research.

2.1 The Central Bank Communication and Economic Conditions

Research on central bank communication has expanded rapidly in recent years, driven by the increasing availability of digitized policy texts and advances in automated text analysis methods. This shift marks a departure from earlier approaches that relied heavily on manual coding, event study designs, or survey based classifications [Ehrmann and Fratzscher, 2009, Blinder et al., 2008, Macklem and Vardy, 2023]. Modern computational tools now make it possible to process large corpora of speeches and statements, though often at the cost of reducing rich linguistic content into a limited set of measurable features. Among these, sentiment-based indicators—typically capturing positive versus negative tone, dissent, or hawkish versus dovish language—have become especially prominent. This line of work spans a wide range of communication types (policy statements, press conferences, speeches) and central banks, and links sentiment to outcomes such as policy uncertainty and market volatility [Baranowski et al., 2021, Tillmann and Walter, 2019, Gu et al., 2022], macroeconomic conditions including unemployment, output, and inflation [Bohl et al., 2023, Clements and Reade, 2020, Eugster and Uhl, 2024], interest rate dynamics [Astuti et al.], and financial market responses [Picault and Renault, 2017, Ehrmann and Talmi, 2020, Schmeling and Wagner, 2023, Hayo et al., 2022, Parle, 2022, Dossani, 2021].

However, the study of ambiguity and linguistic complexity in central bank communication remains comparatively less developed. The dominant proxy for this dimension is readability, which has gained traction due to its intuitive interpretation, its documented influence on financial markets, and its relevance for public understanding of monetary policy. The concern is straightforward: when policy messages are excessively opaque, households may struggle to interpret them, and firms may rely on intermediaries—such as news outlets or analysts—whose interpretations can be subjective, unevenly accessible, or costly. Although the general public historically paid limited attention to central bank communication outside major policy events, recent episodes such as the COVID 19 pandemic and the subsequent inflation surge have heightened public awareness of monetary policy and increased interest in clear communication.

The most widely used readability metric is the Flesch–Kincaid score, which has been shown to rise over time for both the Federal Reserve and the ECB [T. Hughes and Kesting, 2014], prompting the ECB to adopt reforms aimed at simplifying its language. Readability measures have been used to study a range of economic effects, including trading volume and return volatility [Smales and Apergis, 2017], market reactions during periods of unconventional monetary policy [Hayo et al., 2022], and interactions with sentiment indicators [Celler, 2024]. Beyond major central banks, similar analyses have been conducted for smaller or emerging market institutions [Binette and Tchebotarev, 2019, Doan et al., 2023], and a handful of studies compare textual features across countries [Born et al., 2014, Luangaram and Wongwachara, 2017]. However, these contributions remain limited by data availability and, importantly, by their focus not on complexity (or if any, very little), and certainly not utilizing the broader linguistic complexity measures.

Existing work, therefore, leaves open the question of whether more nuanced measures of complexity—beyond readability—carry informational content or predictive power. Our study attempts to address this gap by employing recently developed complexity metrics to evaluate their explanatory value relative to the conventional tools used in the literature. In doing so, we assess whether these richer linguistic features provide additional insight or whether, like many volume-based or readability-based measures, they exhibit limited predictive significance.

On the other hand, understanding the intrinsic characteristics of the texts themselves—how they are constructed and how they evolve over time—is also of considerable interest. In this regard, a part of the literature adopts a reverse approach, regressing textual features (mostly sentiment

variables) on macroeconomic indicators such as real GDP growth, unemployment, and inflation Bohl et al. [2023], often incorporating lag structures to capture temporal dynamics [Hayo and Zahner, 2023]. With respect to complexity, far less is known, although recent work has begun to make meaningful progress. A recent IMF contribution [Silva et al., 2025] uses large language models to classify central bank statements according to semantic clarity and explanatory depth, showing that complexity tends to rise during periods of heightened uncertainty. While their focus is on communicative intent and audience segmentation, Tran (2025) examines linguistic structure—readability, abstractness, informativeness, and coherence—offering a complementary perspective on how complexity responds to economic conditions, particularly during crises.

Building on this line of research, the present study considers similar questions but in the context of individual FOMC members rather than cross-country institutional differences. These two perspectives can yield distinct insights. Unlike national policy statements, which are highly standardized and formal, speeches by individual members of a monetary policy committee reflect personal communication styles and professional backgrounds. In this setting, individual heterogeneity plays a much larger role than institutional history or culture. Moreover, individual preferences are generally easier to analyze than institutional preferences, which can be ambiguous. This helps explain why the literature on FOMC members is relatively rich in studies linking personal backgrounds to communication and behaviour, although very little work examines linguistic complexity. This gap motivates the next section, which reviews the relevant literature and outlines our contribution.

2.2 Individual Heterogeneity in FOMC Members' Communication

Empirical studies of FOMC members have long documented substantial differences across policymakers, drawing on textual data as well as background characteristics and institutional roles. Because the previous section reviewed work that uses FOMC speeches to predict economic outcomes, this section focuses instead on studies that examine how individual attributes shape communication itself.

A large portion of this literature links members' educational or professional backgrounds to their communication styles, particularly in terms of hawkishness and dovishness. For example, Malmendier et al. [2021] show persistent heterogeneity in both communication tone and policy

preferences by combining speech data with detailed biographical information. Numerous other contributions similarly rely on either textual features or background variables to infer the heterogeneous nature of communication within the committee Bennani et al. [2020], Hansen et al. [2019], Acosta [2024], Malmendier et al. [2021], Rae et al. [2022], Bordo and Istrefi [2023]. Evidence from outside the United States also points to the importance of individual leaders: Madeira (2026), for instance, emphasizes the role of the central bank governor rather than the institution as a whole in shaping communication.

Another strand of research highlights how institutional roles—such as voting status or whether a member is a Reserve Bank President or a Board Governor—affect communication strategies [Djourelova et al., 2025, Cannon, 2015, Tillmann, 2011, Claussen et al., 2012]. These findings reinforce the idea that communication is not uniform across policymakers but instead reflects differences in authority, incentives, and institutional positioning.

Economic conditions also play a major role in shaping communication, which is unsurprising given the cyclical nature of monetary policy [Meade and Sheets, 2005]. However, despite the extensive literature on individual heterogeneity in tone, sentiment, and policy preferences, very little is known about whether similar patterns arise in the linguistic complexity or ambiguity of FOMC speeches. This gap motivates our study: we examine whether members' backgrounds can be linked to systematic differences in speech complexity—much like the well-documented relationships between background characteristics and hawkishness, dovishness, or emotional tone—while also accounting for prevailing economic conditions.

3 Data

3.1 FOMC Members' speeches

The textual dataset consists of speeches delivered by all 19 FOMC members, including all Reserve Bank Presidents and Board Governors, up to 7239 number of speeches. The core of the dataset is an extension of Campiglio et al. [2025], which covers member speeches from January 1986 to December 2023. I update this dataset through web scraping of individual Reserve Bank websites and the Federal Reserve's central website, extending coverage through the end of 2025. Although most regional banks maintain consistent formatting and categorization under "remarks"

or "speeches," some entries are less clear, occasionally including conversational events involving multiple participants. In these cases, the Bank President typically dominates the discussion, so the text still largely reflects the member's views and communication style. To maintain consistency with the Campiglio dataset, certain differences are intentionally omitted. For example, YouTube video transcripts (e.g., from the Atlanta Fed), scraped and punctuated using Python tools and labelled as "Vid," are excluded because the original dataset does not contain video-based material. The final conclusion can be tested against the results that include these later on for a robustness check.

Because the scraping process captures footnotes and references, these elements are retained for consistency with the Campiglio dataset. However, I remove abnormally long tokens—such as URLs, unintelligible strings, and other noise introduced by scraping—to reduce distortions that could interfere with a few complexity measures.

A further challenge concerns archival completeness. Many districts have sparse or inconsistent speech records prior to the mid 1990s due to incomplete digitization. After this period, speech frequency becomes more stable. Some districts, such as San Francisco and New York, maintain unusually rich archives extending back to the 1980s. A notable issue arises for the St. Louis Fed: William Poole's speeches (1998–2008) are entirely absent from the Campiglio dataset. To address this gap, I supplement the dataset with materials from the FRASER archive, which contains his full set of speeches during that period. Another distinction concerns the Federal Reserve Bank of New York. Unlike most regional banks, whose archives primarily contain speeches delivered by their Presidents, the New York Fed also publishes remarks from other senior officials who hold influential operational roles, such as the First Vice President or heads of key policy areas. This broader set of speakers reflects the New York Fed's unique position within the Federal Reserve System, particularly its responsibility for implementing monetary policy and conducting market operations. Because these officials often articulate institutional views and policy stances on behalf of the Bank, and because earlier datasets in the literature (including Campiglio et al., 2025) also retain these speeches, all such New York Fed communications are included in the dataset. However, to assess the robustness of this specification, I also estimate all models using an alternative dataset that excludes these additional speakers. The full set of analyses is conducted on both versions of the data. I only report deviations that materially differ from the main specification; in the absence

of such differences, the alternative results are not mentioned at all.

3.2 FOMC Members' Background Information

For the initial descriptive analysis, information on member characteristics is collected from Federal Reserve and regional bank websites, including data for former presidents. The primary objective is to link professional background—particularly educational and career history—to potential differences in communication behaviour, with a focus on complexity. I also encode the specific roles held by Board Governors and Reserve Bank Presidents, distinguishing among chair, vice chair, and other governor positions for the former, and between president and non-president roles for the latter. This distinction is essential because individuals may transition across multiple roles over their careers, and they do not occupy a single position throughout the sample period. An illustrative example is Janet L. Yellen, who served as Chair of the Board of Governors (2014–2018), Vice Chair (2010–2014), President of the Federal Reserve Bank of San Francisco (2004–2010), and Governor (1994–1997). All such transitions are fully encoded. Table 1 lists the Chairs serving between 1985 and 2025 along with their term dates, and Figure 1 provides a Gantt-style overview of role transitions across all three subgroups.

Chair	**Term Start**	**Term End**
Paul Volcker	1979-08-06	1987-08-11
Alan Greenspan	1987-08-11	2006-01-31
Ben Bernanke	2006-02-01	2014-01-31
Janet Yellen	2014-02-03	2018-02-03
Jerome Powell	2018-02-05	Present

Table 1: Federal Reserve Chairs and Their Terms (1985–2025)

For the regression analysis, I draw on a combination of speaker-background datasets, specifically the Federal Reserve Bank Boards of Directors Biographical Database Overview compiled by Peter Conti-Brown and Kaleb Nygaard, together with the dataset provided by Riboni and Ruge-Murcia [2025]. Constructing a complete biographical panel from primary sources alone would require extensive archival work and introduce unnecessary scope for measurement error. Using these established datasets, therefore, ensures consistency, comparability, and transparency in how member characteristics are coded.

These sources allow me to assemble a comprehensive background dataset for FOMC members, including gender, race (white vs. non-white), degree major (economics, law, business/finance, and

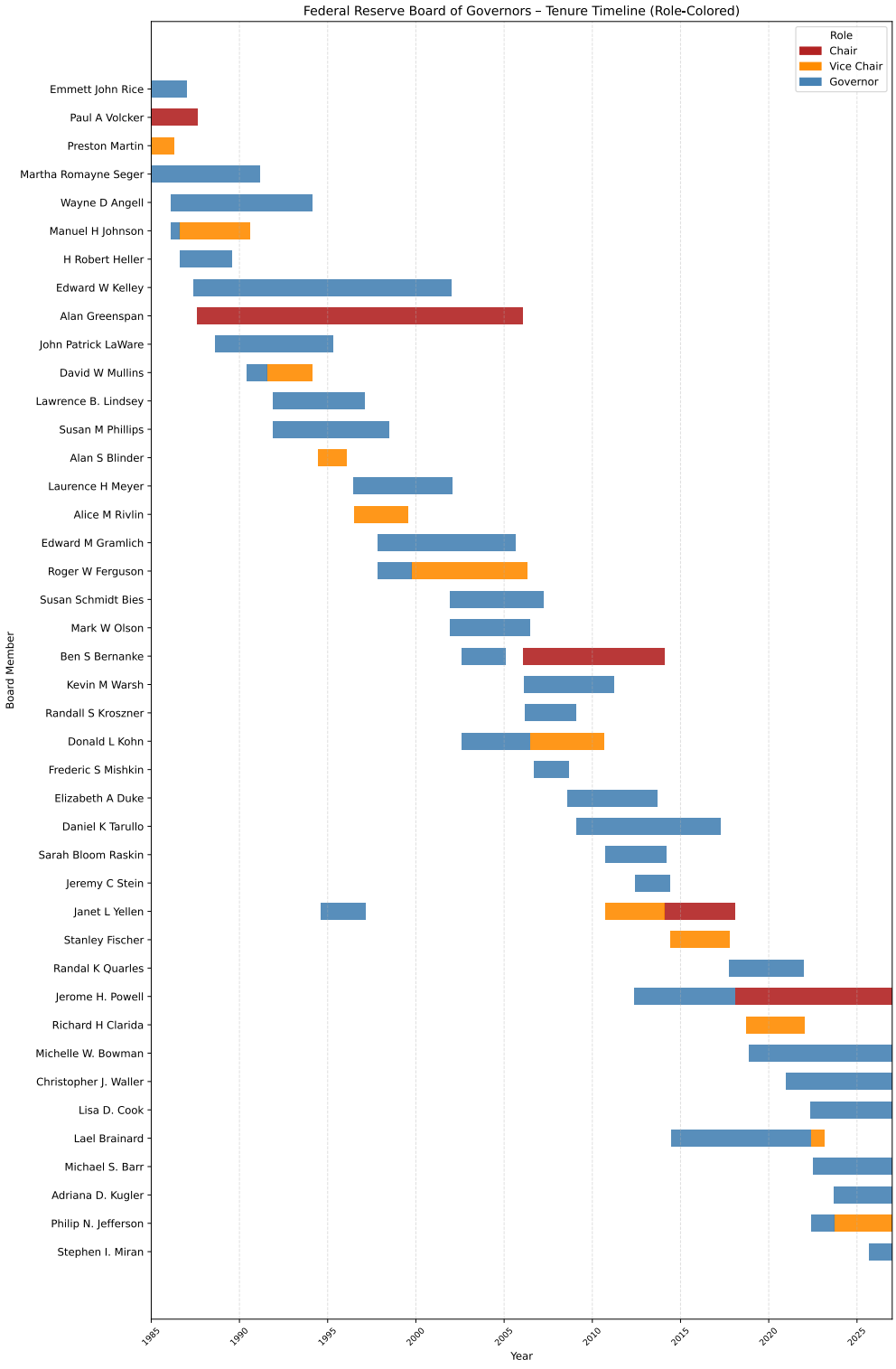

Figure 1: Federal Reserve Board of Governors Timeline

other fields), terminal degree (PhD vs. other), pre-Fed career (economist, academic economist, business/finance, law, and other), saltwater/freshwater classification, age, and years since the start of their term. Degree, major, and institution are taken from the Conti-Brown & Nygaard dataset, using the terminal degree as the reference point. For the pre-Fed career variable, I distinguish between economists working in policy institutions and academic economists, as this distinction is particularly relevant for analyzing communication complexity: differences in training environments may plausibly translate into differences in linguistic style.

Finally, the saltwater/freshwater classification follows the definitions in Riboni and Ruge-Murcia [2025]. This distinction broadly separates economists trained in the "saltwater" tradition—typically associated with coastal universities emphasizing Keynesian and policy-interventionist approaches—from those in the "freshwater" tradition, often linked to inland institutions emphasizing market-clearing models and micro-founded macroeconomics. Including this variable allows me to test whether these intellectual lineages correspond to systematic differences in communication (complexity) behaviour.

3.3 Financial Market Data

To examine how central banks' members' speeches interact with financial markets, I collect daily financial variables spanning the entire length of the dataset, while intraday data only span from January 2010 to December 2023 for reasons explained later. The dataset includes a set of widely used equity, fixed income, currency, volatility, and commodity instruments (intraday data only includes a subset of these). These assets are chosen because they represent the primary channels through which monetary policy communication is expected to transmit to financial markets.

In addition to the asset-price variables, I include a set of timing controls to account for predictable intraday and daily patterns in financial markets. These controls ensure that measured price movements reflect responses to the speeches rather than systematic market regularities. First, I incorporate day-of-the-week indicators, focusing in particular on Monday and Friday effects, as financial markets often display distinct return and volatility patterns at the beginning and end of the trading week.

Second, I include time-of-day controls that distinguish among morning, midday, and afternoon speeches. Intraday liquidity, volatility, and price discovery vary substantially across the trading

session: market activity is typically elevated at the open and close, while midday trading tends to be calmer. Speeches released outside regular trading hours cannot be matched to intraday financial data and are therefore excluded from the sample. As a result, an "outside-trading-hours" dummy would be mechanically zero for all included observations and is omitted to avoid collinearity. These timing controls isolate the portion of intraday financial movements that can plausibly be attributed to the speech rather than to routine intraday trading dynamics.

3.3.1 Daily Data

Daily data is taken from Yahoo Finance using their API pipelines, including the following instruments:

1. **U.S. Equity Market (Spot and Futures)**

   - **S&P 500 Index (SPY, ticker: $\hat{\mathbf{G}}$SPC)** — Represents the broad U.S. equity market and serves as the standard benchmark for aggregate stock-market performance.
   - **E-mini S&P 500 Futures (ES, ticker: ES=F)** — A forward-looking equity instrument that incorporates expectations about future macroeconomic and policy conditions.

   Both are included because they capture complementary dimensions of equity-market reactions: SPY reflects contemporaneous market valuations, while ES incorporates forward-looking expectations and often adjusts more rapidly to news.

2. **U.S. Treasury Market (Rates and Duration)**

   - **SHY — iShares 1–3 Year Treasury Bond ETF**, capturing short-maturity Treasury yields that are highly sensitive to monetary policy expectations.
   - **IEF — iShares 7–10 Year Treasury Bond ETF**, representing intermediate-term Treasury yields.
   - **TLT — iShares 20+ Year Treasury Bond ETF**, reflecting long-term interest rate expectations and duration risk.

   These instruments collectively span the short, intermediate, and long ends of the yield curve, allowing me to assess how communication complexity relates to policy-sensitive interest-rate movements.

3. **U.S. Dollar (Foreign Exchange)**

   - **UUP — Invesco U.S. Dollar Index Bullish Fund**, a tradable proxy for broad U.S. dollar strength. This helps evaluate whether communication complexity correlates with global conditions or shifts in USD demand.

4. **Volatility and Risk Premium**

   - **VXX — iPath S&P 500 VIX Short-Term Futures ETN**, capturing short-horizon market volatility and risk sentiment; commonly interpreted as a measure of risk aversion.
   - **VIX — CBOE Volatility Index**, the standard benchmark for implied equity-market volatility derived from S&P 500 options.

   Both VIX and VXX are included because they measure distinct aspects of market volatility: VIX reflects implied volatility from S&P 500 options, whereas VXX tracks short-term VIX futures and therefore captures tradable, forward-looking risk-premium dynamics.

5. **Safe-Haven / Alternative Asset**

   - **GLD — SPDR Gold Shares**, a highly liquid safe-haven asset that responds to macroeconomic uncertainty and shifts in risk appetite

Together, these instruments provide a comprehensive view of how complexity may influence, or be influenced by, equity valuations, interest rates, currency markets, volatility, and commodity prices.

For each instrument, I begin with the raw daily data (Open, High, Low, Close, and Volume). From these series, I construct a set of standardized financial variables commonly used in empirical asset pricing and event study research.

1. **Returns.** Daily returns measure the immediate market reaction:

   - Simple return: $\text{ret}_t = \frac{P_t - P_{t-1}}{P_{t-1}}$
   - Log return: $\log(P_t) - \log(P_{t-1})$

   These variables capture directional movements in response to complexity.

2. **Volatility Proxies.** To assess whether complexity affects market uncertainty, I construct several volatility measures:

   - Absolute return: $|\text{ret}_t|$
   - Squared return: ret_t^2
   - High–low range (Parkinson estimator): $\text{hl\_range}_t = \log\left(\frac{\text{High}_t}{\text{Low}_t}\right)$

   These proxies capture both realized volatility and intraday price dispersion

3. **Liquidity and Trading Activity.** Trading volume is included as a measure of market attention and liquidity.

4. **Abnormal Returns.** Abnormal returns isolate unusual market movements: abn_ret(5) = return - rolling mean over 5 days; abn_ret(22) = return - rolling mean over 22 days. I construct rolling measures using 5-day and 22-day windows, which correspond to one trading week and one trading month, respectively. These horizons are standard in empirical finance because they capture short-term and medium-term market conditions.

3.3.2 Intraday Data

Before turning to the construction of the intraday dataset, it is important to note that using high-frequency financial data requires precise timestamps for each speech. Unlike central-bank announcements or scheduled press conferences—where release times are fixed and publicly known—FOMC members' speeches are published at irregular intervals. This irregularity is also one reason why the speech corpus cannot be treated as panel data. Moreover, timestamp information is not available in the main textual datasets. To address this, I rely on the timestamped archive provided by Federal Reserve Bank of St. Louis [2024], which covers the period from January 2010 to December 2023. Consequently, the intraday financial data are collected only for this interval. The total number of speeches covered is 1429.

A complication arises because a subset of speeches in the archive is assigned a timestamp of 00:00:00. Since FOMC members do not deliver speeches at midnight, these entries likely reflect missing or automatically generated timestamps rather than actual release times. Such observations cannot be reliably matched to intraday financial data and are therefore excluded. In addition,

speeches delivered outside regular U.S. trading hours cannot be paired with intraday price movements. After applying these restrictions, the final number of usable speech–market observations is approximately 700, with some variations depending on the event window (which I chose 15, 30, 45 and 60 minutes) and the specific financial instrument.

Intraday financial data are obtained from the Bloomberg Terminal, which provides high-frequency, quality-controlled price series widely used in empirical macro-finance research. The construction of the intraday variables mirrors the approach used for the daily data, with the key difference being the sampling frequency: instead of end-of-day observations, I extract 15-minute intervals during regular U.S. trading hours (09:00–16:30 ET). Restricting the sample to trading hours ensures that price movements reflect active market responses rather than illiquid overnight adjustments or stale quotes.

The intraday dataset includes the following instruments: SPY (as a liquid proxy for the S&P 500), SHY, IEF, TLT, and the VIX index. These assets span equities, short-, intermediate-, and long-term Treasury yields, and market-implied volatility, providing a comprehensive view of high-frequency market reactions to FOMC communication.

3.4 Macroeconomic Data

Macroeconomic data are collected on the principle that it is the state of the economy that the speaker observes at the time of the speech. Thus, I am using two different sources of macroeconomic variables, one of which comes from FRED/ALFRED and the other comes from the Green/Tealbook.

3.4.1 Vintage Data from FRED/ALFRED

To incorporate macroeconomic conditions into the analysis, I construct a speech-level dataset in which each observation is matched to the appropriate real-time macroeconomic vintage available on the date the speech was delivered. For each speech date, I retrieve the corresponding vintage series, compute the relevant transformations—such as year-on-year and month-on-month growth rates, inflation rates, quantile indicators, and lagged values—and attach these to the speech record. This procedure is necessary because the speech corpus is not a panel: speeches occur at irregular intervals, and policymakers do not speak on a fixed schedule. Using real-time vintages ensures that the macroeconomic information available to policymakers at the moment of communication is

accurately reflected in the dataset.

The macroeconomic variables include:

- **GDPC1** — Real Gross Domestic Product (quarterly, chained 2017 dollars), the broadest measure of aggregate economic activity.
- **INDPRO** — Industrial Production Index (monthly), a high-frequency indicator of real activity capturing manufacturing output, mining, and utilities.
- **PCEPILFE** — Personal Consumption Expenditures Price Index excluding food and energy ("core PCE"), the Federal Reserve's preferred measure of underlying inflation.
- **CPIAUCSL** — Consumer Price Index for All Urban Consumers (headline CPI), a widely followed measure of consumer-price inflation that includes food and energy.
- **UNRATE** — Civilian Unemployment Rate, a key labour-market indicator reflecting slack and cyclical conditions.
- **MICH** — University of Michigan Consumer Sentiment Index, capturing household expectations and perceptions of economic conditions.

These variables jointly capture real activity, inflation, labor-market conditions, and household sentiment—four dimensions that are central to the Federal Reserve's policy framework and communication strategy.

To capture real economic conditions in a way that aligns with the irregular timing of central-bank speeches, I use industrial production (INDPRO) growth as the primary measure of real activity. Industrial production is available at a monthly frequency, making it substantially more responsive to short-run macroeconomic fluctuations than quarterly GDP. This higher frequency is particularly important in a setting where speeches do not occur at fixed intervals and may cluster around periods of heightened economic attention. Moreover, industrial production is widely used in macro-finance and monetary-policy research as a timely indicator of the business cycle, especially because it reacts quickly to changes in manufacturing output, energy production, and inventory dynamics. These features make it a natural choice for modeling the economic backdrop against which policymakers craft their communication.

For inflation, I focus on headline CPI (CPIAUCSL) in the baseline specification. While core PCE is the Federal Reserve's preferred policy measure, headline CPI plays a distinct role in shaping the public's perception of inflation because it incorporates food and energy prices—components that are highly salient to households and heavily covered in the media. Since the objective of this study is to understand how policymakers adjust the complexity and sentiment of their communication in response to economic conditions, it is reasonable to expect that the Fed may respond differently when inflation pressures are visible to the public. Headline CPI therefore provides a more direct measure of the inflation environment that policymakers anticipate their audience reacting to, especially during periods of volatile consumer prices.

To ensure that the results are not driven by the specific choice of macroeconomic indicators, I conduct a series of robustness checks using alternative measures. First, I replace CPI with core PCE inflation (PCEPILFE), the measure most closely aligned with the Federal Reserve's policy framework, to verify that the results are not sensitive to the inflation concept. Second, I substitute industrial production with real GDP growth (GDPC1) and a bottom-decile GDP-growth indicator, which captures broader and lower-frequency movements in economic activity as well as nonlinear "bad-times" dynamics. Across these alternative specifications, the main findings remain stable, confirming that the relationship between macro-financial conditions and speech characteristics is not an artifact of variable selection but reflects a robust underlying pattern in policymakers' communication behaviour. For this reason, I present the main results using the subset of vintage macroeconomic variables described above.

3.4.2 Greenbook/Tealbook Data

Another important source of macroeconomic information available to FOMC members at the time of their speeches is the staff forecast contained in the Greenbook/Tealbook. These forecasts are prepared for each FOMC meeting—typically eight times per year—providing a frequency that is higher than quarterly data but lower than monthly indicators. For each speech date, I match the speech to the most recently released Greenbook/Tealbook forecast available prior to the speech. This approach mirrors the logic used for real-time vintage data: it approximates the information set that policymakers actually observed when formulating and delivering their speeches and remarks.

From the Greenbook/Tealbook, I extract real GDP growth, inflation, the unemployment rate,

and housing starts. Industrial production growth is excluded because it is highly correlated with real GDP growth in this dataset and, unlike in ALFRED, both series are available at the same forecast frequency; in this context, real GDP growth is the more comprehensive measure of real activity. For each variable, I use the past-period estimate, the current-period assessment, and the staff's forecast for future periods. This is a key advantage of the Greenbook/Tealbook relative to real-time vintage data from FRED: it allows me to capture not only the contemporaneous economic environment but also the forward-looking expectations that may shape how policymakers communicate.

A limitation of this dataset is the release schedule. Greenbook/Tealbook materials are made public with a five-year lag, meaning that the available data currently extend only through the end of 2020. As a result, the sample of speeches that can be matched to staff forecasts is smaller than the sample used for the vintage-data analysis. Nonetheless, the inclusion of staff forecasts provides a valuable complementary perspective, allowing me to examine whether communication complexity responds not only to observed economic conditions but also to the Federal Reserve staff's expectations about the near-term economic outlook.

The data described above form the basis for linking economic conditions to the content and structure of FOMC communication. I now turn to the methodological framework used to characterize the speeches, including the construction of complexity measures and the implementation of the topic-modelling technique.

4 Methodology

This section outlines the methodological framework used to characterize the content and structure of FOMC speeches. I begin by describing the extraction of the text-based complexity measures, including readability, abstractness, informativeness, disunity, and cognitive strain. I then introduce the topic-modelling procedure, detailing the estimation approach, the resulting topic–keyword structure, and the grouped topic categories used in the regression analysis. Finally, I present the three main groups of regression specifications that form the core of the empirical analysis, which examine how the complexity of FOMC communication responds to macroeconomic conditions, with particular emphasis on financial-market variables.

4.1 Complexity Measurements

The complexity measures used in this study follow the methodological framework developed in Tran (2025), which provides the full algorithmic details and examples in monetary policy statement contexts. Here, I summarize the key components and describe how they are adapted to the context of FOMC speeches, which can be slightly different from policy statements. Additionally, I introduce another dimension called "cognitive strain", which will be discussed in more detail after readability, abstractness, informativeness, and disunity.

While the role of complexity in policy statements is often tied to institutional hedging and the clarity of policy transmission, complexity in individual speeches serves a different but equally important economic function. Speeches are discretionary, personally authored communications in which policymakers explain their reasoning, interpret data, and signal their individual policy leanings. In this context, linguistic complexity can reveal how confidently a member interprets economic conditions, how much uncertainty they perceive, and how they choose to frame their arguments to different audiences. More complex speeches may reflect deeper analytical framing, greater caution, or an attempt to justify a nuanced position. Conversely, simpler speeches may indicate stronger conviction, a desire to reach a broader audience, or alignment with the institutional baseline. Because financial markets increasingly parse individual speeches for information about future policy, understanding the complexity of these communications provides insight into how policymakers process information, how they attempt to influence expectations, and how individual heterogeneity shapes the overall communication environment of the FOMC.

4.1.1 Readability

Readability captures the educational effort required to process a text. Following Tran (2025), I compute a broad set of established readability indices—including the Gunning–Fog Index, SMOG, Flesch Reading Ease, Flesch–Kincaid Grade Level, ARI, Coleman–Liau Index, LIX, and RIX. These metrics rely on features such as sentence length, syllable counts, and word complexity. Because these indices are highly correlated, I extract a single composite readability measure using the first principal component from a PCA. Higher values correspond to more difficult, less accessible speech.

4.1.2 Abstractness

Abstractness reflects the degree to which a text relies on intangible or conceptually demanding language. Word-level concreteness scores are obtained from standard psycholinguistic databases (Brysbaert et al., 2013; Scott et al., 2019; Coltheart, 1981). These scores are aggregated into speech-level measures and then inverted so that higher values indicate greater abstractness. As in Tran (2025), multiple variants of the abstractness metric are combined using PCA to form a single composite indicator.

4.1.3 Informativeness

Informativeness is measured using Shannon entropy, which captures the diversity of vocabulary within a speech. Higher entropy indicates a broader distribution of word usage, reflecting richer informational content. Because lexical diversity does not always align with cognitive difficulty, this measure is treated as a distinct dimension of complexity. The entropy values are normalized before being incorporated into the analysis.

4.1.4 Disunity

Disunity (the inverse of coherence) captures the degree of semantic discontinuity across sentences. Sentence-level embeddings are used to compute cosine similarity between adjacent sentences (first-order coherence) and between sentences separated by one intervening sentence (second-order coherence). Lower similarity implies greater semantic distance and therefore higher disunity. As in the other dimensions, the first principal component of these coherence measures is used as the composite disunity indicator.

4.1.5 Cognitive strain

Cognitive strain—often operationalized through dependency distance—is a syntactic complexity measure that captures the distance between each token and its syntactic head or dependent within a sentence. For any given document, this produces a distribution of dependency distances from which both the mean and the standard deviation can be computed. For example, in the sentence "The central bank unexpectedly raised interest rates," the verb "raised" functions as the head of

the clause; "bank" is its subject, modified by "central"; "rates" is its object, modified by "interest"; and "unexpectedly" is an adverbial modifier. Longer dependency distances require readers to hold earlier elements in working memory while processing later ones, thereby increasing cognitive load and reducing reading ease.

A distinctive feature of this measure is that each sentence contains multiple head–dependent pairs, meaning that a single speech yields a distribution rather than a single value. Relying solely on the mean can therefore be misleading when the dispersion of distances varies substantially across documents. To address this, the final cognitive strain measure is defined as the mean plus two standard deviations, which approximates the upper bound of syntactic load within a speech and better reflects the maximum cognitive effort required from the reader. However, the mean variant will still be studied, and the differences, if any, will be documented.

All of the features are normalized due to the process of principal component analysis. This includes single-variable features for better comparison.

4.1.6 Illustrative Examples

To provide intuition for the complexity measures, I include illustrative word clouds comparing speeches in the top and bottom quantiles of abstractness, as well as examples contrasting Reserve Bank Presidents and Board Governors. These visualizations highlight the types of vocabulary associated with different levels of complexity. A more detailed interpretation of these patterns and their relationship to economic and financial conditions is presented in the results section. Figure 2 shows the word clouds for high and low abstractness levels.

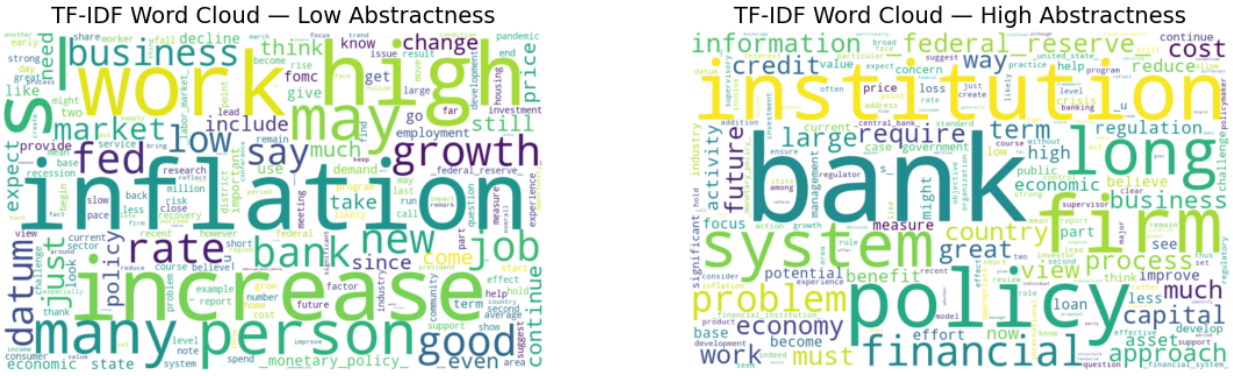

Figure 2: Word clouds for low vs. high abstractness speeches

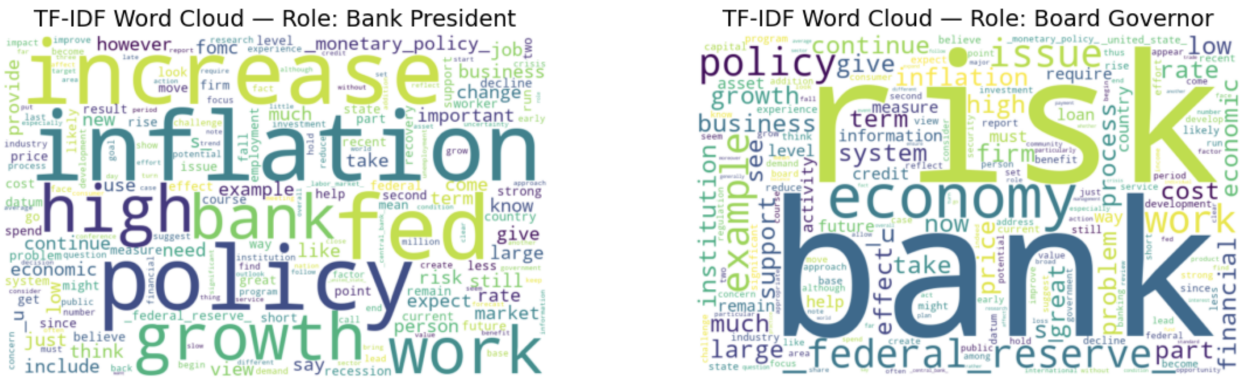

Figure 3: Word clouds for Bank Presidents vs. Board Members

The word clouds provide a useful visual illustration of how the abstractness measure maps onto the vocabulary used in FOMC speeches. As shown in the figures, terms such as system, policy, and institution are more prevalent in speeches with high abstractness, whereas words like inflation, growth, and work appear more frequently in speeches with low abstractness. Although the complexity measures differ in the word clouds' contents, a consistent pattern emerges across them: speeches that emphasize concrete economic conditions—particularly growth and labour-market developments—tend to fall in the lower-complexity range, while speeches that focus on risk, policy frameworks, and systemic considerations tend to fall in the higher-complexity range. This pattern is visible both in the quantile-based wordclouds and in the comparison between Reserve Bank Presidents and Board Governors (Figure 3). The presidents' speeches feature vocabulary tied to real-side economic activity, whereas the governors' speeches more often highlight stability/risk and system themes.

As shown later in the regression analysis, these linguistic distinctions align closely with the behaviour of financial-market variables, particularly measures of volatility and market attention.

4.2 Topic Modelling

To identify the thematic structure of FOMC communication, I apply Latent Dirichlet Allocation (LDA), a widely used probabilistic topic-modelling method. LDA assumes that each document is a mixture of latent topics and that each topic is characterized by a distribution over words. This framework is well-suited for central-bank speeches, which typically blend multiple themes—such as economic outlook, policy strategy, financial stability, and institutional communication—within a

single document. By estimating the underlying topic structure, LDA provides a data-driven way to summarize the content of speeches and to construct topic-level indicators for the empirical analysis.

Before estimating the model, I preprocess the corpus following standard text-analysis procedures. I begin by tokenizing the speeches and applying lemmatization to reduce inflected forms to their base representation. I then identify multi-word expressions (n-grams) that capture meaningful economic or institutional phrases. Specifically, I retain four-grams and three-grams that appear at least 50 times in the corpus, and I include two-grams that exceed the 50th percentile of the likelihood-ratio statistic, which helps filter out spurious collocations while preserving economically relevant expressions. After removing conventional stop words and special characters, I construct the document–term matrix using a refined stop-word list tailored to central-bank communication. This list excludes numbers, high-frequency boilerplate terms, generic modal or filler words, formatting artifacts, non-economic speech conventions, magnitude words (e.g., percent), weak connectors (e.g., include), and time expressions that do not meaningfully distinguish topics.

To determine the appropriate number of topics, I estimate LDA models with k = 10, 20, 30, 40, 50 topics and evaluate them using standard coherence and exclusivity diagnostics. A model with k = 40 topics provides the best balance between granularity and interpretability: it is sufficiently detailed to capture the diverse themes present in FOMC communication without producing overly fragmented or redundant topics. The resulting topic structure forms the basis for the grouped topic categories used in the regression analysis.

The full set of 40 LDA topics and their associated keywords is reported in Appendix 7.4. These topics span a wide range of themes—including monetary policy, financial stability, macroeconomic outlook, housing and credit, international spillovers, technology, fiscal policy, and pandemic-related communication—and provide a granular representation of the thematic structure of FOMC speeches.

To incorporate thematic information into the empirical analysis, I aggregate the 40 raw LDA topics into a set of economically interpretable groups. These grouped topics capture major dimensions of FOMC communication—including monetary policy, financial stability, macroeconomic outlook, housing and credit, international spillovers, technology, fiscal policy, and pandemic-related themes—while excluding low-coherence or metadata topics. The full topic–keyword list and the detailed grouping procedure are provided in Appendix 7.4.

To characterize how linguistic complexity varies across topics, I compute topic-level averages

of each complexity measure using the topic proportions as weights. This approach is standard in the topic-modelling literature because LDA assigns each document a distribution over topics rather than a single categorical label. Topic proportions therefore represent the intensity with which each topic is expressed in a document, and weighting by these proportions provides a coherent way to aggregate document-level attributes to the topic level. This method preserves the soft-assignment structure of LDA and avoids arbitrary thresholding or hard classification of documents into topics (Blei, Ng & Jordan, 2003; Blei, 2012). Similar weighting strategies are used in empirical applications where researchers relate document-level outcomes to topic intensities, including studies of central bank communication and political text (e.g., Hansen & McMahon, 2016; Gentzkow, Kelly & Taddy, 2019). Table 2 shows the correlations.

topic	Community development	Fed operations payments	Financial stability regulation	International economics	Monetary policy inflation	Real economy productivity
Abstract	-0.444	-0.015	1.152	0.437	-0.414	-0.525
Informativeness	-0.372	0.116	0.172	-0.033	-0.014	-0.102
Readability	-0.174	0.029	0.611	-0.067	-0.144	-0.259
Disunity	-0.229	0.215	0.433	0.118	0.065	-0.314
Strain	-0.088	0.119	0.253	0.017	0.032	-0.283
Strain (Upper)	-0.057	0.412	0.118	-0.173	0.039	-0.175

Table 2: Correlation between Complexity Measures and Grouped Topics

The table shows that different themes carry distinct linguistic signatures rather than reflecting a single, uniform notion of "complexity." Topics related to growth and development—such as community development, monetary policy and inflation (including pandemic topics), and real-economy productivity—tend to exhibit much lower complexity, especially in terms of abstractness. These themes rely on concrete, widely understood economic concepts—such as jobs, wages, prices, and business conditions—which naturally require fewer abstract or technical terms. Because they are often directed at broad public audiences, the language tends to be more narrative and accessible, resulting in systematically lower levels of conceptual and syntactic complexity.

By contrast, more technical themes—financial stability and regulation, Federal Reserve operations and payments, and international economic conditions—display substantially higher complexity. In particular, financial stability and international topics score especially high on abstractness, with the former correlated positively (relatively high) across all complexity measures. Fed operations and payments do not exhibit high abstractness, but they show pronounced spikes in the upper

bound of cognitive strain, indicating that certain sentences or constructions are especially difficult to process. This pattern likely reflects the hybrid nature of operational communication: much of the content might be intentionally simplified for public audiences, keeping abstractness low, yet key passages still require precise technical descriptions, resulting in spikes in cognitive strain, elevating its upper bound relatively more.

Taken together, the table suggests that complexity is multidimensional: topics differ not only in how much complexity they contain but also in the type of complexity they express—whether conceptual, informational, or syntactic.

4.3 Regressions

The empirical analysis consists of two groups of regression specifications, each designed to address a different purpose. The first set examines whether financial-market conditions in the days preceding a speech have predictive power for the complexity of that speech, even after controlling for a rich set of speaker-level and document-level characteristics. The second set reverses the timing and tests whether the complexity of today's speech helps predict next-day financial-market movements, using daily financial data. Additionally, we also run regressions that follow the same logic as the second but employ intraday financial data to assess whether the high-frequency response of markets mirrors the dynamics observed at the daily level, consistent with the existing literature on monetary-policy communication.

4.3.1 Financial Data as Predictors of Speech Complexity

For the first group of regressions, we run the following OLS specification:

$$LC_{i,t} = \beta_0 + \sum_{k=1}^{4} \beta_k Fin_{t-k} + \gamma X_{i,t} + \alpha_i + \epsilon_{i,t} \tag{1}$$

where $LC_{i,t}$ denotes the complexity measure of speech i delivered on day t. This includes 6 different measurements, including abstractness, readability, informativeness, disunity, cognitive strain and upper bound of cognitive strain. Fin_{t-k} represents financial-market variables observed k days prior to the speech. The vector $X_{i,t}$ contains control variables that vary at lower frequencies, including monthly macroeconomic indicators (INDPRO growth, CPI inflation, Michigan sentiment,

unemployment rate, and their lags) as well as topic proportions. Because these controls are not observed daily, each speech is matched to the most recently available monthly or quarterly value as of date t (see the Data section for details). For robustness, I also consider alternative macroeconomic controls: (1) replacing CPI with PCE inflation, (2) replacing industrial production with real GDP growth, and (3) using Greenbook/Tealbook forecasts. Only the Greenbook/Tealbook specifications are reported in the appendix; any differences arising from (1) or (2) are noted in the text. Speaker fixed effects α_i absorb time-invariant heterogeneity across individuals, implemented using 150 dummy variables corresponding to unique speakers. All regressions are estimated using ordinary least squares with Newey–West heteroskedasticity- and autocorrelation-consistent (HAC) standard errors, employing a Bartlett kernel with five lags. This correction is applied uniformly across all OLS specifications in the paper.

Given the large number of financial variables, I do not interpret individual coefficients directly. Since the objective is to assess whether financial conditions preceding a speech have predictive power for its complexity, I conduct an F-test of the joint null hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ against the alternative that at least one lag of the financial variable predicts speech complexity. Because this test does not reveal the direction of the effect, I also test the linear restriction $H_0 : \beta_1 + \beta_2 + \beta_3 + \beta_4 = 0$ using one-sided alternatives to assess whether the cumulative effect is positive or negative.

In addition, I examine the coefficients on the control variables, which provide valuable insights beyond the financial predictors. To do so, I highlight one representative regression specification using SPY volume as the financial variable. Several alternative specifications were estimated, and since the main conclusions remain unchanged, only the SPY-volume results are reported.

Finally, I conduct an ANOVA-style variance-decomposition analysis to quantify the contribution of different groups of variables—macroeconomic controls, speaker background, financial variables, and topic proportions—to the explained variation in speech complexity. For this purpose, the financial variable group includes all non-overlapping measures for each instrument. The macroeconomic group is again split into two sets: one using FRED vintage data (including all available controls) and one using Greenbook/Tealbook forecasts, which is used as a robustness check. The speaker-background group includes all variables described in the Data section, supplemented by either institutional dummies or a role indicator distinguishing Reserve Bank Presidents from

Board Governors.

4.3.2 Speech Complexity as Predictors of Financial Data

For the second group of regressions, we run the following OLS specification:

$$Fin_{t+1} = \beta_0 + \sum_{k=1}^{6} \beta_k LC_{i,t}^k + \gamma X_{i,t} + \epsilon_{i,t} \tag{2}$$

where $LC_{i,t}^k$ denotes the complexity measure k for speech i delivered on day t, with k = 1,...,6 corresponding to abstractness, readability, informativeness, disunity, cognitive strain, and the upper bound of cognitive strain. Fin_{t+1} represents financial-market variables observed a day after the speech. The vector $X_{i,t}$ contains a set of controls similar to those used in the first group of regressions, including monthly macroeconomic indicators and topic proportions. In addition, I include day-of-week indicators (Monday and Friday) and a dummy for whether the speaker is the Chair. Because the control structure mirrors that of the first group, the same robustness checks apply; only the third alternative specification is reported in the appendix.

As in the previous section, I conduct hypothesis tests on the coefficients associated with the linguistic-complexity measures to assess whether speech complexity has predictive power for next-day financial outcomes, and, when relevant, to determine the direction of the effect.

The intraday version is very similar except that the outcome variables required more steps to compute, the control variable sets contain an additional set of variables—the time of the day dummies, and the number of observations is drastically smaller:

$$\Delta Fin_{t,window} = \beta_0 + \sum_{k=1}^{6} \beta_k LC_{i,t}^k + \gamma X_{i,t} + \epsilon_{i,t} \tag{3}$$

For the construction of the intraday outcome variables, I consider a set of event windows w = 15, 30, 45, 60 measured in minutes relative to the start time of each speech. For each window, I compute two types of high-frequency financial responses. The first measures the change in the financial variable from the moment the speech begins to w minutes after the start, that is $Fin_{t,start+w} - Fin_{t,start}$. The second measures the change from w minutes before the speech to w minutes after the speech, that is $Fin_{t,start+w} - Fin_{t,start-w}$. The former captures the immediate post-speech reaction, while

the latter captures the net movement around the speech event, controlling for pre-speech trends. These two constructions allow the analysis to be robust between pure speech-induced movements and broader intraday dynamics, with the former as the main result.

5 Results

5.1 Descriptive Analysis

Given the size and richness of the dataset—nearly four decades of speeches from every FOMC participant—it would be neither tractable nor conceptually meaningful to analyze communication patterns speaker-by-speaker. Instead, the structure of the Federal Reserve itself provides a natural hierarchy for organizing the results. I begin by comparing Board members and Reserve Bank presidents, two groups that differ systematically in their institutional roles, audiences, and communication incentives, making them a theoretically relevant starting point for understanding variation in linguistic complexity. More reasoning will be provided in the later subsection. I then narrow the focus to the Board of Governors, where internal role differentiation—particularly between the Chair, Vice Chair, and other Governors—creates well-known asymmetries in influence and communication strategy. Finally, I turn to the twelve Reserve Banks individually, recognizing that each president operates within a district and institutional context that may shape their speech style. This tiered approach allows the analysis to move from broad structural contrasts to finer individual distinctions more systematically, highlighting the dimensions of variation that are most relevant for understanding how monetary policymakers communicate.

5.1.1 Board Members

Differences in communication style between Federal Reserve Board members and Reserve Bank presidents are well documented in the monetary policy literature, even though these distinctions have rarely been examined through the lens of linguistic complexity. Board members operate within a highly centralized institutional environment, speak more frequently on behalf of the Committee as a whole, and face stronger political and media scrutiny. These constraints tend to shape their rhetoric toward caution, consensus-building, and strategic ambiguity. In contrast, Reserve Bank presidents often tailor their communication to regional audiences and enjoy greater freedom to ex-

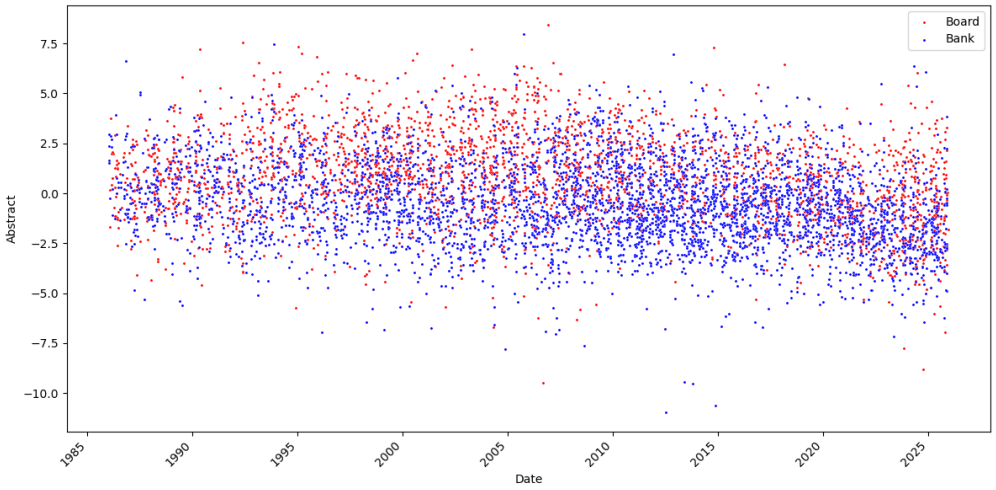

Figure 4: Board Members versus Bank Presidents: Abstractness

press individual views, a pattern highlighted in studies of dissent behavior, speech tone, and regional emphasis in policy communication. These established differences provide a natural motivation for examining whether they also extend to deeper linguistic features—such as abstractness, readability difficulty, informativeness, disunity, and cognitive strain—which remain largely unexplored in the existing literature.

Using scatter plots of each complexity measure over time, we can visually assess whether systematic differences emerge across groups or institutional roles. As a first step, we compare Board members with Reserve Bank presidents. Overall, Board members exhibit slightly higher levels of complexity across four of the five measures—abstractness, readability difficulty, disunity, and cognitive strain—while for informativeness, there are no clear differences. These differences are particularly pronounced before 2000. For illustration, Figure 4 and Figure 5 present the time-series scatter plots for abstractness and readability, respectively; the remaining measures are reported in the Appendix. In these figures, red dots denote Board governors and blue dots denote Reserve Bank presidents.

All figures show values fluctuating around zero by construction of the normalization, with no strong upward or downward trend over the sample period, with the exception of abstractness. Although subtle, abstractness exhibits a mild downward drift after 2020, which may reflect changes in communication style during the post-pandemic period. While one could speculate that the higher complexity observed among Board members relative to Reserve Bank presidents reflects

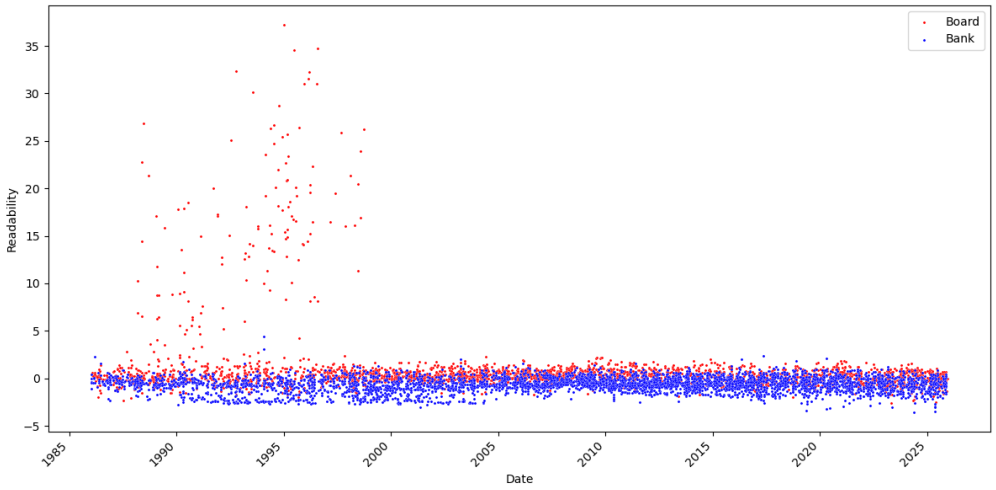

Figure 5: Board Members versus Bank Presidents: Readability

differences in institutional duties and incentives, many other factors may also contribute. For this reason, it is useful to examine variation within the Board itself—specifically, whether occupying the role of Chair or Vice Chair leads to systematically different linguistic behavior compared with serving as a Governor. This is particularly relevant given the unusually high levels of complexity observed before 2000, which may be linked to broader economic conditions or, more plausibly, to the communication style of the Chair during that period.

To analyze these internal distinctions, I present the same scatterplots as before, but with the red (Board governors) and blue (Reserve Bank presidents) points lightly faded to emphasize the black points corresponding to the subgroup of interest (e.g., the Chair). This visualization makes it clear that the most extreme values of complexity—across abstractness, readability difficulty, disunity, and cognitive strain—are overwhelmingly associated with Chair speeches, especially prior to 2000, with a modest decline after 2020. Vice Chairs also tend to exhibit above-average complexity and often track the Chair's patterns, although they deliver far fewer speeches. The remaining Governors, reflecting a heterogeneous group of individuals over time, display no consistent trend and naturally exhibit a wider dispersion. A representative example of these patterns can be seen in the abstractness plots shown in Figures 6, 7, and 8.

For both the Chair and the Vice Chair, speeches generally lie at the upper end of the complexity distribution, even when compared only within the Board member group. Abstractness shows a slight decline after 2020 for both roles, although the pattern is not particularly pronounced. Other

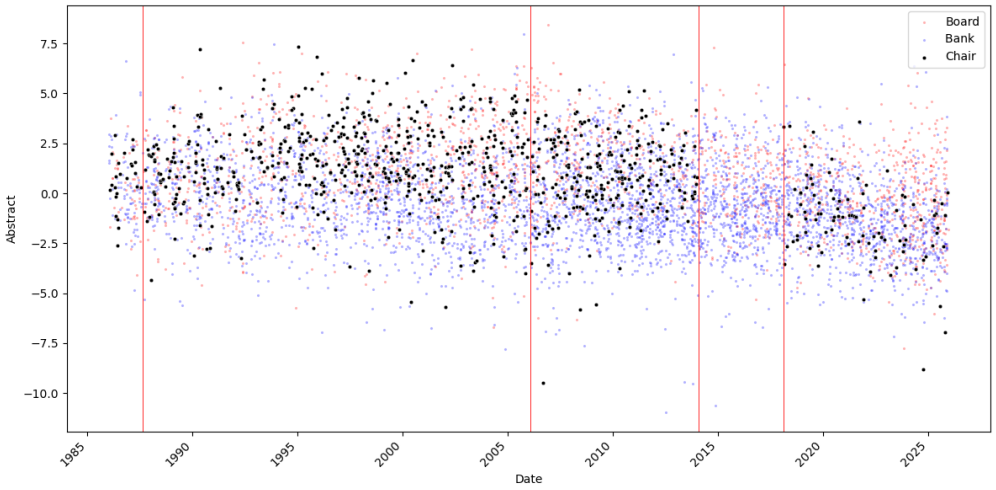

Figure 6: Board Members - Chair: Abstractness

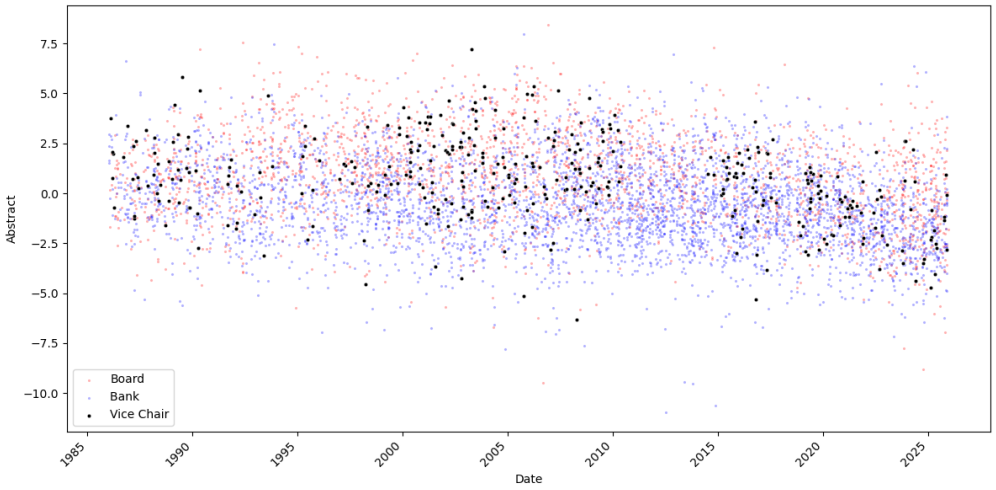

Figure 7: Board Members - Vice Chair: Abstractness

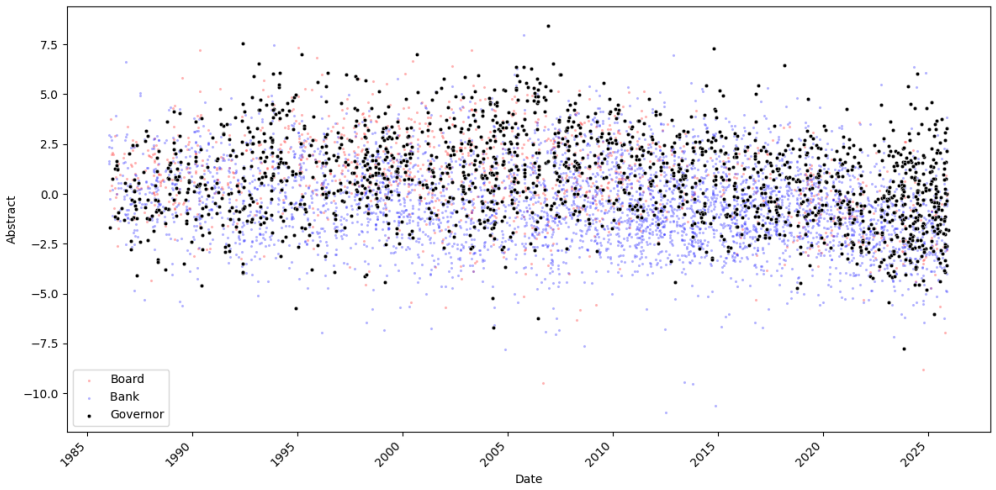

Figure 8: Board Members - other Governors: Abstractness

Governors exhibit a wider dispersion, reflecting the heterogeneity of multiple individuals over time, and their abstractness also appears to drift downward very slightly after 2020. For the remaining complexity measures, the analysis focuses primarily on the Chair. The Governor group is too heterogeneous to support meaningful interpretation, and the Vice Chair delivers far fewer speeches, making it difficult to draw reliable conclusions. Moreover, the Vice Chair's patterns tend to mirror those of the Chair, but without the extreme values observed before 2000.

Before turning to the features that display the most striking behavior, it is helpful to contextualize the Chairs' communication patterns. According to Table 1, with the transitions also marked in the figures by vertical red lines, the period prior to 2000 corresponds to the tenure of Chairman Greenspan, which aligns closely with the elevated complexity observed in the data. Greenspan is widely associated with strategic ambiguity—a deliberate rhetorical approach intended to avoid committing to precise policy paths and to prevent markets from overreacting to his statements. His early years, particularly from the late 1980s through the 1990s, are frequently described in the literature as the era of "Fed-speak," characterized by dense, abstract, and syntactically intricate communication. After 2000, however, the decline in complexity coincides with a broader institutional shift toward greater transparency in monetary policy communication. During the late 1990s and early 2000s, the FOMC began issuing more explicit statements, clarifying policy rationales, and standardizing communication practices—reforms that Greenspan gradually supported despite his earlier resistance. These changes, together with the growing academic consensus that transparency enhances policy effectiveness, likely contributed to the measurable simplification of Greenspan's speeches in the 2000s.

To illustrate these patterns more clearly—and to show that the narrative described for abstractness becomes even more pronounced for readability difficulty, disunity, and cognitive strain—I focus on the disunity scatterplots in Figures 9 and 10; the remaining measures (for the Chair only) are reported in the Appendix.

As the figures show, the Chair exhibits markedly elevated complexity from the beginning of Chairman Greenspan's tenure through roughly 2000, in sharp contrast to the relatively stable range displayed by the other Governors. After 2000, Greenspan's speeches become noticeably easier to parse, with a further decline toward the end of his term, although this pattern is less visible when using the upper-bound cognitive strain measure, which remains comparatively stable. A modest

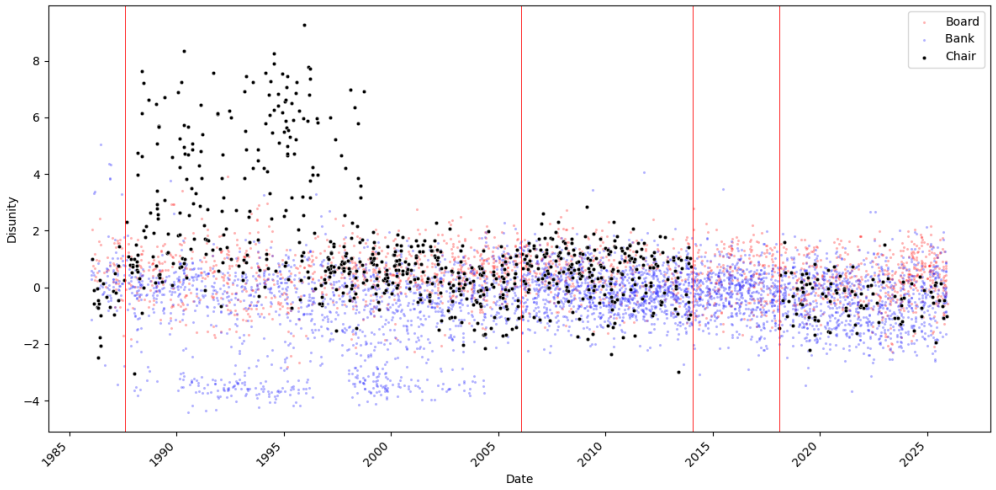

Figure 9: Board Members - Chair: Disunity

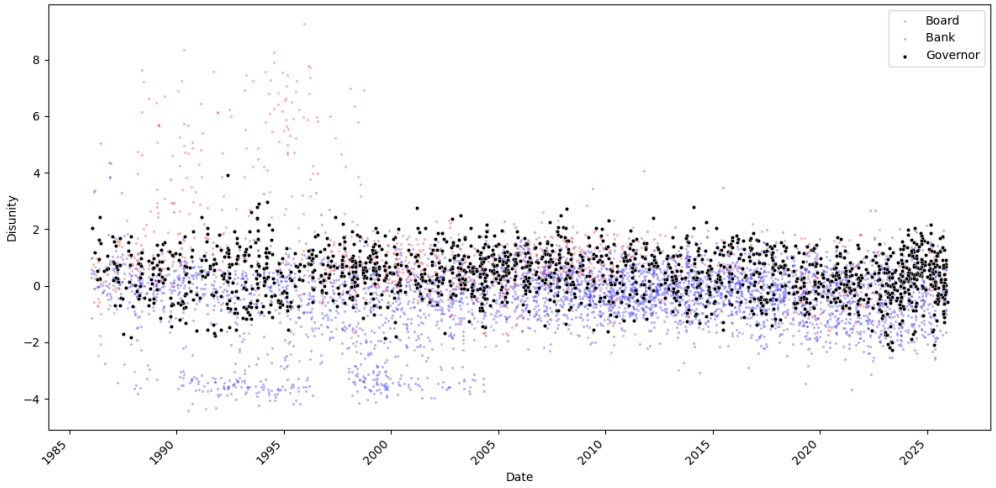

Figure 10: Board Members - other Governors: Disunity

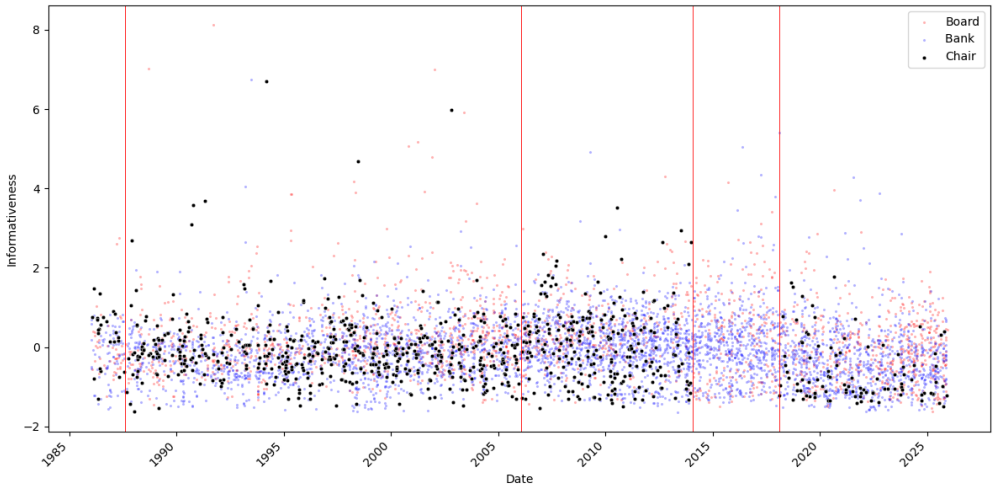

Figure 11: Board Members - Chair: Informativeness

decline in complexity also appears after 2019–2020, during Chairman Powell's tenure, which may reflect both differences in communication style across Chairs and the broader shift in rhetorical tone during the post-pandemic period.

The final feature, informativeness, behaves quite differently from the other measures. Figure 11 presents the Chair's informativeness scores, with the corresponding plots for the Vice Chair and Governors included in the Appendix.

Across all three groups, a common pattern emerges: informativeness rises gradually until around 2008, after which it becomes highly polarized, with speeches clustering at either very high or very low levels. This polarization appears regardless of institutional role. After 2020, the trend converges downward on average, but with a noticeably wider dispersion, suggesting greater heterogeneity in how policymakers conveyed information during the post-pandemic period.

In conclusion, we observe a clear and meaningful difference in linguistic complexity between Board Governors and Reserve Bank presidents, with Chairman Greenspan's pre-2000 speeches standing out as a prime example of the strategic ambiguity and limited transparency characteristic of that era. This pattern provides some validation for the complexity measures used, as they successfully capture the well-known period of "Fed-speak." By contrast, informativeness appears less well-suited to identifying this historical shift, as it does not fully reflect the heightened complexity of Greenspan's early tenure. Within the Board itself, there is consistent evidence that the Chair—and, to a lesser extent, the Vice Chair—exhibits higher-than-average complexity relative to

other Governors, who show no systematic trend and span a much wider range of linguistic behavior. With this Board-level structure established, the analysis now turns to the twelve Reserve Banks individually, where regional mandates, institutional differences, and leadership turnover provide a richer setting to assess whether background characteristics are correlated with speech complexity.

5.1.2 The 12 District Bank Presidents

The detailed analysis of each Regional Reserve Bank will be listed in the appendix 7.3. I will just summarize the main observations below.

Archival Completeness Issues: a consistent feature across many Reserve Banks is that speech archives only become reliably populated in the mid-1990s or early 2000s. This affects Banks such as Boston, Philadelphia, Cleveland, Chicago, Kansas City, Dallas, and Minneapolis, where earlier presidents appear to have delivered fewer speeches simply because records were not systematically preserved or digitized. As a result, early-period speech counts should not be interpreted as reflecting true communication behavior. Only New York and San Francisco maintain extensive archives reaching back to the 1980s, allowing for more robust long-run comparisons. These archival differences matter because they shape the apparent evolution of communication complexity: early presidents often appear less active simply due to missing data. Thus, interpreting speech frequency early on should be done with caution.

System-Wide Shifts After Crisis such as the Global Financial Crisis (GFC) and the Pandemic: After the GFC, nearly every district shows an increase in speech frequency and possibly a modest rise in complexity, reflecting the need for more detailed explanations of unconventional monetary policy and evolving policy frameworks. Presidents across the System engaged more actively with public audiences, and speeches became more analytically dense. In contrast, the pandemic period brought a sharp decline in abstractness and a shift toward more practical, operational communication. This pattern appears in New York, Boston, Cleveland, Chicago, Kansas City, San Francisco, and others, regardless of the president's academic or professional background. This is only apparent in abstractness, not other measurements (if there are any, it is fairly uncommon).

The Role of Professional Background in Shaping Communication Complexity: across districts, a president's professional background is one of the strongest predictors of communication complexity. Presidents with academic macroeconomic training—such as Plosser (Philadelphia),

Lacker (Richmond), Mester (Cleveland), Yellen (San Francisco), Kocherlakota (Minneapolis), and Logan (Dallas)—consistently produce speeches with higher abstractness, greater disunity, broader vocabulary, and more intricate structure. Their communication, therefore, often incorporates themes such as theoretical frameworks, long-run policy considerations, and analytical reasoning, which is more likely to boost the complexity of speeches. By contrast, presidents with business, finance, or managerial backgrounds—such as Harker (Philadelphia), Pianalto (Cleveland), Barkin (Richmond), Moskow (Chicago), Lockhart (Atlanta), Kaplan (Dallas), and George (Kansas City)—tend to produce more concrete, narrative-driven speeches with lower complexity. Their communication is then likely to be shaped by applied experience, operational concerns, and managerial knowledge. A third group consists of presidents with highly public-facing or rhetorical styles, such as McTeer and Fisher in Dallas, whose speeches are accessible, cohesive, and often stylistically distinctive. These patterns demonstrate that professional identity and career trajectory play a central role in shaping communication style across the Federal Reserve System.

To conclude, there is evidence that shows that communication complexity across the Federal Reserve System is shaped by a combination of professional background, institutional role, and macroeconomic environment, with significant crises and individual expertise playing a crucial role in the members' communication complexity.

5.2 Regressions Results

5.2.1 Financial Data as Predictors of Speech Complexity

The first group of regression results are shown in Tables 3

The most striking pattern is that the predictive power for speech complexity comes overwhelmingly from volatility-based measures—specifically, absolute returns, high–low price ranges, and market-attention proxies such as trading volume. In contrast, simple returns and abnormal returns (over both 5-day and 22-day windows) exhibit very limited explanatory power for linguistic behavior. Among the few exceptions, short-term Treasury-bond (SHY) returns influence informativeness, while the U.S. dollar index (UUP), short-term volatility futures (VXX), and option-implied volatility (VIX) affect cognitive strain and its upper bound, though only at weak significance levels. For abnormal returns, the signals are scattered around the 10% level and do not form a coherent

Category	Instrument	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Return	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	**	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	0
	UUP	0	0	0	0	*	0
	VXX	0	0	0	0	0	**
	VIX	0	0	0	0	*	0
	GLD	0	0	0	0	0	0
Absolute Return	SPY	0	0	***	***	***	***
	ES	0	0	0	**	0	**
	SHY	0	0	0	0	**	0
	IEF	**	0	0	0	0	0
	TLT	***	0	**	0	0	0
	UUP	0	0	0	0	0	0
	VXX	**	**	0	0	0	*
	VIX	0	0	0	0	0	0
	GLD	0	0	***	*	*	0
High-low Range	SPY	***	0	***	***	***	***
	ES	**	0	0	0	0	0
	SHY	0	0	0	0	0	0
	IEF	**	0	**	0	0	0
	TLT	**	0	***	0	0	0
	UUP	*	0	0	0	0	0
	VXX	0	0	0	0	0	0
	VIX	0	*	**	0	0	***
	GLD	**	0	0	***	0	0
Volume	SPY	**	***	***	***	***	0
	ES	0	0	0	*	0	***
	SHY	***	***	0	0	0	0
	IEF	***	***	**	***	0	**
	TLT	***	***	0	***	0	***
	UUP	**	0	0	0	***	0
	VXX	0	0	**	0	0	0
	GLD	*	0	0	0	0	0
Abnormal Return (5 days)	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	**	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	0
	UUP	0	0	0	0	0	0
	VXX	**	0	0	0	0	0
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0
Abnormal Return (22 days)	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	*	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	*
	UUP	0	0	*	0	*	0
	VXX	0	0	0	0	0	0
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0

Table 3: Financial Data as Predictors of Speech Complexity

Notes: This table reports joint-significance tests of $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ for each asset–complexity pair in regression 1. Each cell presents the corresponding Wald test statistic and p-value. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include speaker fixed effects, macroeconomic controls (vintage ALFRED), and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 3346 to 6468

pattern.

Between the two volatility measures, absolute returns and the high–low price range display

broadly similar levels of explanatory power, though for different types of assets. Both measures show strong to moderate predictive power for broad U.S. equity markets, medium- and long-term Treasury bonds (IEF, TLT), and the gold market (GLD). However, absolute returns additionally predict complexity through short-term Treasury bonds and volatility-futures products, whereas the high–low range does so only through option-implied volatility. This suggests that policymakers adjust their communication in response to broad, macro-relevant volatility that affects both short-term interest rates and volatility-linked instruments. By contrast, the high–low range predicts complexity only through implied volatility, indicating that intraday turbulence matters for communication only when it is sufficiently severe to be incorporated into option prices. This pattern implies that policymakers respond primarily to economically meaningful volatility rather than to transitory intraday noise.

Regarding which aspects of complexity respond most strongly to volatility measures, for broad equity markets, the effects appear across nearly all dimensions except informativeness. Indeed, informativeness is the least predictable measure, showing relevance only from implied-volatility and volatility-futures instruments. For the remaining asset classes, the most responsive dimensions tend to be abstractness and readability, with occasional contributions to cognitive strain and disunity from non-Treasury and non-equity markets.

For trading volume, the strongest predictive power appears in the equity market (excluding equity-index futures) and across the Treasury-bond complex (excluding short-term Treasuries), where explanatory power reaches the 1% level for nearly all aspects of textual complexity. Although weaker, market attention in equity-index futures also predicts disunity (10% level) and the upper bound of cognitive strain (1% level). For short-term Treasuries, trading volume predicts abstractness and informativeness at the 1% level. Other asset classes show scattered predictive signals, but none form a noticeable or consistent pattern.

When examining the sign-hypothesis results in the Table 22, several patterns emerge. As in the main specification, most predictive power comes from volatility-based measures and trading volume. However, the explanatory power of simple returns and abnormal returns is not entirely absent. These measures are particularly informative for disunity and cognitive strain (both average and upper bound), suggesting that policymakers may not react to isolated one-day fluctuations but instead adjust their communication in response to the cumulative pattern of recent market

movements.

In terms of sign, higher volatility and trading-volume levels in the broad U.S. equity market (SPY) predict lower values across almost all dimensions of speech complexity, while higher equity returns predict only lower upper-bound cognitive strain, and even then with limited evidence. This pattern is consistent with policymakers adopting clearer and more reassuring language during periods of system-wide turbulence. By contrast, a different pattern emerges for other asset classes—including Treasury bonds (SHY, IEF, TLT), implied-volatility indices, volatility-futures products, gold, and equity-index futures (ES). Higher levels across these markets tend to predict lower abstractness and informativeness but higher readability, disunity, and cognitive strain (especially at the upper bound). In other words, policymakers' language becomes more fragmented with occasional pockets of syntactic difficulty but more concrete and repetitive when uncertainty arises through hedging and risk-transfer channels rather than through broad equity-market conditions.

The robustness checks using the Greenbook/Tealbook dataset yield results broadly consistent with the main specification, with one notable difference: the influence of volatility and trading volume in the U.S. Treasury market becomes substantially weaker. Specifically, the predictive power of absolute returns diminishes, while the high–low range and volume measures remain present but at somewhat reduced strength. This may suggest that Greenbook/Tealbook information plays an important role for FOMC members when preparing speeches—capturing aspects of the policy environment that vintage data from FRED may not fully reflect.

5.2.2 Investigate the Relationship between Speech Complexity and Control Variables

In the first part of the analysis, I treat speaker identity as a fixed effect and show that financial-market conditions—particularly trading volume in SPY—are strongly associated with variation in speech complexity. In the second part, I shift the focus from eliminating speaker heterogeneity to explicitly modelling it. By replacing individual fixed effects with observable speaker-background characteristics, I examine whether differences in professional experience, policy roles, or institutional positions systematically shape communication style. Importantly, I retain SPY volume and the macroeconomic controls in these regressions to ensure that the estimated background effects are not confounded by contemporaneous financial or macroeconomic conditions. This allows me to interpret the results as reflecting genuine differences in communication behaviour across speakers,

rather than differences in the environments in which they speak. However, when other assets and measures are used, the results retain their core conclusions with variation in significance level rather than signs, and even if they do, it is very limited.

Variable	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Gender	-0.32***	-0.22***	0.01	-0.07*	-0.15***	0.08**
	(0.08)	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)
Race	0.27*	0.06	0.74***	0.47***	0.28***	0.56***
	(0.14)	(0.06)	(0.12)	(0.08)	(0.06)	(0.05)
Major: Law	0.29*	0.09	0.21**	-0.14*	0.09	0.15**
	(0.16)	(0.07)	(0.10)	(0.08)	(0.06)	(0.06)
Major: Business	-0.39***	-0.03	0.54***	0.29***	0.20***	0.26***
	(0.15)	(0.07)	(0.08)	(0.08)	(0.06)	(0.06)
Major: Finance	-1.89***	-0.57**	-0.50***	0.32**	0.10	-0.13
	(0.37)	(0.27)	(0.15)	(0.14)	(0.11)	(0.24)
Major: Other	-0.37***	0.00	-0.28***	-0.12**	0.07*	-0.21**
	(0.10)	(0.10)	(0.05)	(0.05)	(0.04)	(0.08)
Not PhD	-0.15*	-0.11*	-0.37***	-0.19***	-0.27***	-0.38***
	(0.09)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Pre-Fed: Academic	0.39***	0.17***	-0.08**	0.10***	0.10***	-0.04
	(0.08)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
Pre-Fed: Finance	0.22**	0.01	0.19***	0.39***	0.26***	-0.12*
	(0.10)	(0.09)	(0.05)	(0.05)	(0.04)	(0.07)
Pre-Fed: Business	-1.16***	0.52***	-0.33	-0.02	0.36***	0.21**
	(0.25)	(0.11)	(0.20)	(0.12)	(0.09)	(0.11)
Pre-Fed: Law	-0.46**	-0.14	0.12	0.44***	0.45***	-0.04
	(0.19)	(0.10)	(0.12)	(0.11)	(0.08)	(0.09)
Pre-Fed: Other	0.96***	0.06	2.80***	1.36***	0.96***	0.57***
	(0.13)	(0.07)	(0.25)	(0.09)	(0.07)	(0.07)
Freshwater	0.21**	0.19***	0.13***	0.29***	0.19***	-0.14***
	(0.10)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)
Other Economist	-0.12	-0.25***	0.82***	0.53***	0.42***	-0.14***
	(0.10)	(0.04)	(0.08)	(0.05)	(0.04)	(0.04)
Age	-0.01	0.01***	0.04***	0.02***	0.02***	0.01***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Terms Time	0.04***	-0.02***	-0.07***	-0.05***	-0.03***	-0.01*
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Is Bank Pres	-0.57***	-0.22***	-0.39***	-0.44***	-0.35***	-0.44***
	(0.07)	(0.05)	(0.05)	(0.04)	(0.03)	(0.04)
N	5726	5726	5726	5726	5726	5726
R^2_{adj}	0.297	0.105	0.234	0.337	0.305	0.225

Table 4: Financial Data as Predictors of Speech Complexity: Panel A Speaker background

Notes: This table reports coefficient estimates from regression 1 using SPY trading volume as the financial variable. Each column corresponds to a different measure of textual complexity, and each row lists the associated coefficient estimate and Newey–West HAC standard error (Bartlett kernel, 5 lags) for the variable indicated on the left. All specifications include speakers' backgrounds with role dummies, macroeconomic controls, and topic-proportion controls. The number of observations and R^2 for each specification are reported at the bottom of the table. Panel A and Panel B present the results in two parts due to the table size.

The regression results provide a much sharper view of how speaker-level characteristics correlate with linguistic complexity. Both gender and race have significant predictive power: male members generally speak in less complex language than female members, while white speakers tend to use more complex language than non-white speakers.

Comparing speakers by their academic major reveals systematic differences in linguistic style

Variable	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
gINDPRO	4.81	-3.84*	2.94	3.01	1.96	-1.69
	(4.11)	(2.01)	(4.90)	(2.26)	(1.80)	(2.12)
gINDPRO lag1	2.92	0.33	3.54	0.72	1.04	1.90*
	(2.51)	(1.12)	(2.47)	(1.30)	(1.02)	(1.05)
Infl	4.28	0.12	-11.88*	-7.42*	-3.86	-5.27*
	(7.09)	(2.62)	(6.18)	(3.92)	(2.82)	(3.00)
Infl lag1	-6.15	-2.11	5.38	1.25	1.06	0.18
	(6.47)	(2.43)	(5.68)	(3.69)	(2.60)	(2.66)
MICH	-0.17**	0.01	-0.03	0.03	0.02	0.07**
	(0.08)	(0.03)	(0.09)	(0.05)	(0.03)	(0.04)
MICH lag1	-0.07	-0.04	0.14	0.07	0.02	0.01
	(0.09)	(0.04)	(0.09)	(0.05)	(0.03)	(0.04)
Unemp	0.06	-0.09***	0.06	0.01	0.01	-0.03
	(0.07)	(0.02)	(0.06)	(0.03)	(0.02)	(0.02)
Unemp lag1	0.01	0.09***	0.03	0.00	0.01	0.04*
	(0.07)	(0.03)	(0.06)	(0.03)	(0.02)	(0.03)
Topic: Community Dev	-1.00***	-0.75***	0.90*	1.38***	0.38***	0.50***
	(0.30)	(0.14)	(0.48)	(0.19)	(0.14)	(0.15)
Topic: MonPol–Inflation	-0.26	0.34***	1.10**	2.21***	0.82***	0.74***
	(0.26)	(0.12)	(0.47)	(0.17)	(0.13)	(0.12)
Topic: Real Econ–Prod	-1.55***	0.21*	-0.26	0.97***	-0.22	0.23*
	(0.26)	(0.12)	(0.49)	(0.19)	(0.14)	(0.12)
Topic: Fed Ops–Payments	-0.20	0.81***	0.37	2.07***	0.48***	1.99***
	(0.36)	(0.19)	(0.53)	(0.22)	(0.16)	(0.26)
Topic: FinStab–Reg	2.25***	0.60***	1.71***	2.47***	0.87***	0.62***
	(0.28)	(0.15)	(0.47)	(0.18)	(0.13)	(0.12)
Topic: International Econ	1.92***	0.49***	0.38	2.70***	0.94***	0.50***
	(0.33)	(0.15)	(0.52)	(0.22)	(0.16)	(0.17)
N	5726	5726	5726	5726	5726	5726
R^2_{adj}	0.297	0.105	0.234	0.337	0.305	0.225

Table 5: Financial Data as Predictors of Speech Complexity: Panel B macroeconomic controls and topics

Notes: This table reports coefficient estimates from regression 1 using SPY trading volume as the financial variable. Each column corresponds to a different measure of textual complexity, and each row lists the associated coefficient estimate and Newey–West HAC standard error (Bartlett kernel, 5 lags) for the variable indicated on the left. All specifications include speakers' backgrounds with role dummies, macroeconomic controls, and topic-proportion controls. The number of observations and R^2 for each specification are reported at the bottom of the table. Panel A and Panel B present the results in two parts due to the table size.

relative to those trained in economics. Individuals with a background in law tend to produce speeches that are more abstract, more syntactically demanding (higher readability scores), and more difficult in their most complex passages (higher upper-bound cognitive strain), although their overall coherence is slightly better. This pattern likely reflects legal training, which emphasizes conceptual framing and complex sentence structures but also values internal logical consistency.

Speakers with business degrees show a different profile: they use less abstract language but exhibit higher readability, greater disunity, and higher cognitive strain across the board compared with those holding economics degrees. This combination suggests a communication style that is concrete but structurally fragmented—possibly reflecting a focus on practical, operational details rather than conceptual exposition, leading to more abrupt transitions and heavier syntactic load.

Similarly, those with finance degrees tend to use less abstract and less informative language, yet they display higher disunity. This may reflect a communication style shaped by financial-market discourse: concise, data-driven, and concrete, but often organized around multiple scenarios or risk factors, which can disrupt narrative flow.

Degree level also matters. Non-PhD speakers consistently exhibit lower complexity across all measures. This is unsurprising: doctoral training emphasizes theoretical framing, methodological precision, and dense argumentation, all of which naturally translate into more abstract, more information-rich, and more syntactically complex speech. Those without such training tend to communicate in a more accessible, less technical manner.

Pre-Fed professional experience is another strong predictor of linguistic style, and its effects do not always align with those of academic major. Compared with individuals whose prior careers were in general economics roles, former academics (professors or researchers) produce speeches that are more complex across nearly all dimensions. Their language is more abstract, more informative, more syntactically demanding, and more cognitively taxing, consistent with academic norms of exposition and argumentation. The only exception is readability, which is slightly lower (i.e., easier), though the evidence is limited.

Individuals with prior careers in finance also exhibit higher complexity throughout, with strong significance across nearly all measures except the cognitive strain upper bound. This pattern stands in contrast to the linguistic style associated with finance majors—except for disunity—and highlights the differences between academic training and professional experience in finance. Finance professionals routinely explain uncertainty, risk channels, and market mechanisms, which could consistently increase abstraction, syntactic difficulty, and cognitive strain in speaking styles.

Former business professionals show patterns similar to those with business degrees: lower abstractness but higher informativeness and higher cognitive strain (both average and upper bound). Interestingly, the magnitude of the abstractness effect is two to three times larger than the effects on informativeness and cognitive strain. This suggests that business-trained speakers rely heavily on concrete, operational language but still introduce enough detail and qualification to increase processing difficulty. Finally, speakers with legal backgrounds display lower abstractness but higher disunity and higher average cognitive strain.

Several additional speaker characteristics also exhibit statistically significant effects on speech

complexity. Compared with "saltwater" economists, "freshwater" economists tend to produce more complex speeches across nearly all dimensions, with the exception of upper-bound cognitive strain, which is lower. This pattern may reflect differences in intellectual training between saltwater and freshwater graduates.

Age and tenure in office display an interesting contrast. Older policymakers tend to produce more complex speeches across all dimensions except abstractness, suggesting that experience may lead to richer, more nuanced communication. In contrast, the number of years served in the term is associated with lower complexity across most measures, while abstractness increases over time. One interpretation is that as policymakers become more familiar with the institutional environment, they streamline their communication but rely more heavily on conceptual framing, perhaps because they grow more comfortable articulating high-level narratives rather than detailed technical explanations.

Institutional role matters, unsurprisingly, as shown in the descriptive section. Reserve Bank presidents generally produce substantially less complex speeches than Board Governors, and this difference remains robust even when Chairman Greenspan is excluded (see Appendix 7.7). All effects are significant at the 1% level. When replacing role indicators with institutional dummies, the results remain broadly similar: most Reserve Banks exhibit lower complexity than the Board of Governors. This likely reflects the different institutional mandates and communication constraints faced by the two groups. Board Governors—especially those involved in Washington-based policymaking—operate under stricter scrutiny and must articulate national policy positions, which may require more detailed, abstract, and syntactically demanding communication. Reserve Bank presidents, by contrast, often speak to regional audiences and may have greater flexibility to tailor their remarks, resulting in simpler and more accessible language. However, this pattern does not apply uniformly across all Banks or all complexity measures (see Appendix **??**, Table 24)

Macroeconomic conditions do not exhibit strong predictive power for speech complexity. Increases in industrial production growth and inflation generally have negative effects on complexity, suggesting that during periods of economic expansion, speeches tend to be less complex. However, the evidence is limited and appears only at the 10% significance level. The unemployment rate shows a mixed pattern: higher current-month unemployment is associated with lower informativeness, whereas the previous month's unemployment rate has a positive effect on informativeness.

Topic usage, by contrast, shows markedly different results from the simple correlations presented earlier in the methodology section, where no additional controls were included. Technical topics related to financial risks and international economic conditions remain strongly positively associated with complexity across all measures, with significance at the 1% level. The topic related to Federal Reserve operations and payments now predicts not only higher upper-bound cognitive strain but also higher average cognitive strain, higher informativeness, and greater disunity. The real-economy productivity topic remains relatively concrete, yet it is associated with higher disunity (all at the 1% level), which is reasonable given that speeches on productivity often address uncertainty and structural challenges. This is also reflected—though more weakly—in higher informativeness and upper-bound cognitive strain (both at the 10% level).

Community development topics continue to exhibit lower abstractness and lower informativeness, but they are now also associated with higher disunity and higher cognitive strain (both average and upper-bound), all at the 1% level. Finally, the topic showing the most substantial change once controls are included is monetary policy and inflation. This topic is now associated with higher complexity across nearly all measures except abstractness, indicating that once other factors are accounted for, monetary policy communication becomes substantially more demanding in terms of linguistic structure and cognitive load.

Taken together, these results show that speech complexity is shaped significantly by who is speaking and what they are speaking. Although broader economic indicators such as industrial production, inflation, and unemployment provide evidence of a down-trending economy correlated with higher complexity, these have limited explanatory power. Technical and risk-related topics consistently elevate complexity, reflecting the cognitive and linguistic demands of communicating uncertainty, financial mechanisms, and policy transmission channels. At the same time, speaker backgrounds also play a central role: academic training, degree level, professional experience, and intellectual tradition each leave distinct linguistic signatures. Age and tenure further shape communication in opposite ways, while institutional roles—particularly the contrast between Board Governors and Reserve Bank presidents—highlight the influence of organizational context and communication mandates. Taken together, these findings suggest that the speech complexity of FOMC members arises primarily from who is speaking and what they are speaking about, rather than from real-time economic conditions. The next section, to some extent, confirms this.

5.2.3 Speech Complexity Composition: ANOVA Analysis

Table 6 reveals a pattern consistent with the earlier findings: most of the explanatory power for variation in speech complexity comes from speaker characteristics and, even more prominently, from the topical category of the speech, rather than from financial variables or macroeconomic conditions.

Equity	Group	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
SPY	speaker	5.15%	5.20%	8.75%	19.30%	21.22%	11.59%
	finance	0.75%	0.14%	0.3%	0.15%	0.18%	0.27%
	macro	0.73%	0.46%	0.32%	0.21%	0.17%	0.45%
	topics	3.50%	1.86%	0.56%	11.28%	1.07%	3.43%
ES	speaker	7.27%	10.39%	20.90%	15.03%	28.94%	17.70%
	finance	0.46%	0.9%	0.15%	0.31%	0.28%	0.73%
	macro	1.43%	0.93%	0.50%	0.29%	0.65%	0.99%
	topics	3.91%	2.05%	38.05%	25.08%	2.89%	6.37%
SHY	speaker	6.76%	11.28%	19.38%	12.49%	25.77%	17.62%
	finance	1.1%	0.8%	0.41%	0.42%	0.59%	0.83%
	macro	1.35%	0.70%	0.61%	0.35%	0.86%	1.22%
	topics	4.08%	2.13%	39.80%	27.56%	3.31%	6.63%
IEF	speaker	7.29%	11.07%	19.80%	12.19%	25.58%	17.78%
	finance	1.5%	0.52%	0.45%	0.75%	0.39%	0.62%
	macro	1.07%	0.61%	0.65%	0.38%	0.83%	0.90%
	topics	3.92%	2.14%	39.52%	27.47%	3.29%	6.53%
TLT	speaker	7.00%	11.23%	19.37%	12.03%	25.43%	18.21%
	finance	1.3%	0.66%	0.23%	0.38%	0.14%	0.76%
	macro	1.38%	0.63%	0.55%	0.29%	0.67%	0.90%
	topics	4.07%	2.16%	40.07%	27.73%	3.33%	6.30%
UUP	speaker	4.80%	8.60%	14.64%	7.85%	19.67%	13.42%
	finance	0.86%	0.73%	0.29%	0.41%	0.48%	0.8%
	macro	0.82%	0.73%	0.49%	0.35%	0.93%	0.79%
	topics	5.77%	2.43%	49.35%	34.15%	6.30%	8.37%
VXX	speaker	19.30%	16.51%	9.20%	6.88%	13.25%	3.36%
	finance	3.4%	3.2%	0.82%	0.72%	1.9%	2.3%
	macro	3.08%	1.38%	1.81%	0.54%	3.22%	2.03%
	topics	4.02%	3.28%	60.00%	35.89%	10.34%	11.22%
VIX	speaker	6.38%	6.07%	11.29%	21.05%	24.05%	13.17%
	finance	1.4%	0.75%	0.39%	0.097%	0.22%	0.79%
	macro	1.18%	0.53%	0.22%	0.14%	0.11%	0.62%
	topics	3.62%	2.15%	1.07%	12.74%	1.43%	4.65%
GLD	speaker	6.69%	10.27%	17.40%	9.61%	21.09%	16.76%
	finance	0.7%	0.23%	0.45%	0.72%	0.53%	0.11%
	macro	1.23%	0.69%	0.68%	0.35%	0.84%	1.23%
	topics	5.11%	2.49%	43.06%	30.54%	5.12%	7.26%

Table 6: Variance decomposition of linguistic complexity by speaker backgrounds, financial variables, macro variables and topic variables

This table reports the ANOVA decomposition for regression 1. Each row corresponds to a group of explanatory variables, and the columns report the associated sum of squares and share of explained variation for each complexity measure. The decomposition is based on the OLS specification with speaker-background variables, macroeconomic controls, financial variables, and topic-proportion controls. For the financial variable group, the decomposition aggregates all non-overlapping measures of each asset listed on the right-hand side.

However, the magnitude of explanatory power differs substantially across complexity measures and across asset specifications. For SPY and VIX, speaker backgrounds consistently explain the largest share of variation across all dimensions of textual complexity—approximately 5% for

abstractness and informativeness, around 10% for readability and upper-bound cognitive strain, and roughly 20% or more for disunity and average cognitive strain. The remaining three groups contribute much smaller shares, with the exception of the topic group for disunity, which reaches 11–12%.

For the remaining assets, speaker backgrounds continue to explain the most variation among the four groups for abstractness (5–7%, rising to 19.3% for VXX), informativeness (5–11%, rising to 16.51% for VXX), average cognitive strain (20–28%), and upper-bound cognitive strain (11–17%, except for VXX, where topics become the dominant contributor). In contrast, topic categories explain the largest share of variation in readability and disunity, accounting for 40–60% of readability variation and 25–34% of disunity variation. Financial variables and macroeconomic conditions contribute very little across all specifications, generally explaining less than 1% to 4% of the variation.

Finally, across all four groups, the highest explanatory power consistently appears for disunity, readability, and cognitive strain, while abstractness shows the smallest share of explained variation. This suggests that structural and syntactic dimensions of speech—rather than vocabulary-level abstraction—are the most sensitive to differences in speaker identity, topical content, and financial or macroeconomic context.

When the Greenbook/Tealbook data are used in place of the ALFRED vintage series (see Appendix **??** in Table 25), the core results remain largely unchanged. The explanatory power of macroeconomic variables increases slightly in a few specifications—occasionally reaching as high as 8% for VXX—but generally remains very small, typically below 1% as in the baseline analysis. This confirms that the main findings are not sensitive to the choice of macroeconomic data source and that macroeconomic conditions continue to play only a limited role in explaining variation in speech complexity.

5.2.4 Speech Complexity as Predictors of Financial Data

Turning to the second group of regressions, presented in Table 7, we find that volatility measures—such as absolute returns and high–low ranges—as well as trading volume for the equity market (SPY and ES) and for all three Treasury-bond ETFs are significantly influenced by speech complexity, with effects detectable at the 1% level except for two instances involving ES. For the medium-term Treasury ETF (IEF), returns also respond to speech complexity, although only at the 10% signif-

icance level. The high–low ranges of VXX and VIX (which has no volume series) likewise react meaningfully to complexity. Finally, both the trading volume of GLD and its high–low range are affected by the complexity of the speech. The core results remain after replacing ALFRED vintage data with Greenbook/Tealbook macroeconomic data (see Appendix 27).

Asset	Return	Absolute Return	High-low Range	Volume	Abnormal Return (5 days)	Abnormal Return (22 days)
SPY	0	***	***	***	0	0
ES	0	0	0	***	0	0
SHY	0	***	***	***	0	0
IEF	*	***	**	***	0	0
TLT	0	***	***	***	0	0
UUP	0	0	**	0	0	0
VXX	0	0	*	***	0	0
VIX	0	0	***		0	0
GLD	0	0	0	***	0	0

Table 7: Speech Complexity as Predictors of Financial Data: return, volatility, and abnormal return measures across assets

Notes: This table reports joint-significance tests of all complexity measures $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ for each asset measure in regression 2. Each cell presents the corresponding Wald test statistic and p-value. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include macroeconomic controls (vintage ALFRED) and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 3346 to 6468.

The sign-hypothesis results in Table 26 in the appendix show that most of these predictive relationships carry negative signs, with the exception of SPY and VXX trading volume. This indicates that higher speech complexity generally dampens volatility and market attention, effectively cooling financial reactions. One possible interpretation is that more complex communication may signal caution or strategic ambiguity, prompting markets to respond more conservatively.

Table 8 reports both significance and sign hypotheses for the intraday regressions.

Similar to the daily results, only volatility measures and trading volume—not returns—are statistically explained by speech complexity. However, the timing of the response varies across assets. For SPY, both volatility measures react beginning at the 45-minute window and beyond, while trading volume responds at the 30-minute mark but not within 15 minutes. For all three Treasury-bond ETFs (SHY, IEF, TLT), volatility reacts within 15 minutes, with longer-maturity bonds showing effects that persist beyond that window. Trading activity responds later, around 30 minutes, similar to SPY. Finally, VIX reacts only through its high–low range at the 30-minute horizon, and only at the 5% significance level.

	Return				Absolute Return				High-low Range				Volume			
	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m
SPY	0	0	0	0	0	0	***	0	0	0	*	**	0	***	**	***
SHY	0	0	*	0	*	0	0	0	**	0	0	0	0	**	0	**
IEF	0	0	0	0	***	0	0	*	**	0	**	0	0	***	**	***
TLT	0	0	0	0	*	0	0	0	*	0	**	0	0	**	0	***
VIX	0	0	0	0	0	0	0	*	0	**	0	0				

Table 8: Speech Complexity as Predictors of Financial Data with Intraday Data: Post-speech Reaction

Notes: This table reports joint-significance tests of all complexity measures $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ for each asset measure in regression 3. Each column corresponds to a specific intraday window for a given asset measure, and each cell presents the corresponding Wald test statistic and p-value. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include macroeconomic controls (vintage ALFRED) and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 787 to 847.

However, the sign-hypothesis results tell a different story for the intraday regressions (see Appendix Table 28). Specifically, although trading volume shows consistently negative reactions beginning at the 30-minute window and onward—mirroring the daily-frequency results—the volatility measures behave differently. The only significant intraday volatility responses occur for medium- and long-term Treasury bonds within the first 15 minutes after the speech, and these effects are positive, in contrast to the negative signs observed in the daily regressions. One possible interpretation is that the immediate market response captures short-lived uncertainty or information-processing noise, which later dissipates or reverses once markets absorb the full content of the speech. In other words, complexity may initially generate brief volatility spikes in duration-sensitive assets, even though the overall daily effect is stabilizing.

For the symmetric intraday specification around the speech window (see Appendix Table 29), the only robust reactions come from trading-volume measures, with magnitudes similar to those in the main specification. SPY and IEF additionally show significant responses within the 15-minute windows before and after the speech, rather than at the 30-minute mark. By contrast, most reactions in absolute returns and high–low ranges are not statistically significant in this symmetric setup.

Taken together, the second group of regressions shows that speech complexity systematically affects market activity, primarily through volatility and trading volume rather than returns. At the daily frequency, higher complexity generally dampens volatility and reduces trading activity, suggesting that complex communication may act as a stabilizing force by tempering immediate

market reactions. Intraday results reveal a more nuanced pattern: duration-sensitive assets such as medium- and long-term Treasury bonds exhibit brief volatility increases within the first 15 minutes after complex speeches, consistent with short-run information-processing uncertainty, while the broader equity market (SPY) reacts more slowly and primarily through trading volume rather than volatility. These short-lived intraday spikes ultimately reverse at the daily horizon, aligning with the overall calming effect observed in the baseline results.

Together with the findings from the first group of regressions, the evidence suggests a two-way interaction in which market turbulence encourages policymakers to adopt more complex communication, and complex communication in turn moderates market responses once information is fully absorbed.

6 Conclusions

This paper set out to examine three central questions regarding the determinants and consequences of linguistic complexity in FOMC communication. The empirical results provide clear and internally consistent answers to each.

First, policymakers do adjust the complexity of their communication in response to financial-market conditions observed prior to the speech. Across a wide range of assets, volatility-based measures—absolute returns, high–low price ranges, and trading volume—are the strongest and most reliable predictors of speech complexity, not returns or abnormal returns. These effects are broad-based for equity markets and Treasury bonds and extend to volatility-linked instruments such as VIX and VXX. Importantly, the direction of adjustment varies by market: turbulence in the broad equity market (SPY) is associated with simpler, more reassuring language, whereas volatility in other asset classes—Treasuries, gold, volatility futures—predicts more fragmented, syntactically demanding, and cognitively taxing communication. These patterns suggest that policymakers respond not to isolated price movements but to the broader structure of market uncertainty, adjusting their communication style in ways that reflect the nature and source of financial stress.

Second, individual policymakers differ systematically in their communication styles, and these differences are strongly linked to observable characteristics. Academic background, degree level, pre-Fed professional experience, institutional role, and intellectual tradition all leave distinct lin-

guistic signatures. Legal training is associated with abstract and syntactically complex language; business and finance majors produce more concrete but more fragmented speech; PhD holders consistently speak in more complex terms; former academics and finance professionals exhibit the highest overall complexity; and Reserve Bank presidents speak in substantially simpler language than Board Governors. Topic choice further amplifies these differences: technical, risk-related, and policy-intensive topics generate the highest levels of complexity, while community development and real-economy topics tend to be more concrete but structurally fragmented. ANOVA results confirm that speaker characteristics and topic categories explain far more variation in complexity than financial or macroeconomic conditions, underscoring the central role of speaker identity and content in shaping communication.

Third, the complexity of FOMC communication influences financial-market behaviour in the period following the speech. Complexity systematically predicts volatility and trading volume—but not returns—across equities, Treasuries, gold, and volatility indices. At the daily frequency, the effects are predominantly negative: more complex speeches dampen volatility and reduce trading activity, suggesting that complexity may function as a stabilizing force by tempering market reactions. Intraday results reveal a more nuanced dynamic: duration-sensitive assets exhibit brief volatility increases within the first 15 minutes after complex speeches, consistent with short-run information-processing uncertainty, while broader equity markets respond more slowly and primarily through trading volume. These short-lived intraday spikes ultimately reverse by the daily horizon, aligning with the overall calming effect observed in the baseline results.

Taken together, the evidence points to a two-way interaction between financial markets and FOMC communication. Market turbulence—especially outside the broad equity index—encourages policymakers to adopt more complex, qualified, and structurally demanding language. In turn, complex communication moderates market responses once information is fully absorbed, reducing volatility and trading intensity at the daily horizon. This dynamic highlights the dual role of FOMC speeches as both reflections of prevailing financial conditions and instruments that shape market behaviour.

While the evidence reveals a clear two-way interaction between financial markets and FOMC communication, these patterns should be interpreted as empirical associations rather than structural causal mechanisms. The analysis relies on observational data and reduced-form specifications,

which limit the ability to isolate exogenous variation in either market conditions or speech complexity. Moreover, the complexity measures, though comprehensive, capture only observable textual features and may omit deeper semantic or strategic dimensions of communication. Future research could develop a more rigorous theoretical framework or leverage quasi-experimental designs to strengthen causal inference. In addition, this paper focuses on financial markets as the primary audience; extending the analysis to the general public—who may be less equipped to process highly technical or complex language—would shed light on how communication effectiveness varies across audiences, not just across speakers. Finally, applying the framework to other central banks could illuminate how communication strategies evolve across institutional contexts and policy environments.

References

Miguel Acosta. The perceived causes of monetary policy surprises. *American Economic Review*, 114(11):3493–3534, 2024.

Philippe Andrade and Filippo Ferroni. Delphic and odyssean monetary policy shocks: Evidence from the euro area. *Journal of Monetary Economics*, 117:816–832, January 2021. ISSN 0304-3932. doi: 10.1016/j.jmoneco.2020.06.002. URL https://www.sciencedirect.com/science/article/pii/S0304393220300787.

Valerio Astuti, Alessio Ciarlone, and Alberto Coco. The role of central bank communication in inflation-targeting eastern european emerging economies. URL https://papers.ssrn.com/abstract=4295466.

Spandan Banerjee, Rajendra Paramanik, Rounak Sil, and Unninarayanan Kurup. When all speak, should we listen? A cross-country analysis of disagreement in policymaking and its implications. *Economic Notes*, 53, February 2024. doi: 10.1111/ecno.12234.

Paweł Baranowski, Wirginia Doryń, Tomasz Łyziak, and Ewa Stanisławska. Words and deeds in managing expectations: Empirical evidence from an inflation targeting economy. *Economic Modelling*, 95:49–67, February 2021. ISSN 0264-9993. doi: 10.1016/j.econmod.2020.12.003. URL https://www.sciencedirect.com/science/article/pii/S0264999320312645.

Martin Baumgärtner and Jens Klose. Why central banks announcing liquidity injections is more effective than forward guidance. *International Finance*, 24(2):236–256, 2021. ISSN 1468-2362. doi: 10.1111/infi.12389. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/infi.12389. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/infi.12389.

Gillinder Bedi, Facundo Carrillo, Guillermo A. Cecchi, Diego Fernández Slezak, Mariano Sigman, Natália B. Mota, Sidarta Ribeiro, Daniel C. Javitt, Mauro Copelli, and Cheryl M. Corcoran. Automated analysis of free speech predicts psychosis onset in high-risk youths. 1(1):1–7, 2015. ISSN 2334-265X. doi: 10.1038/npjschz.2015.30. URL https://www.nature.com/articles/npjschz201530. Publisher: Nature Publishing Group.

Hamza Bennani and Matthias Neuenkirch. The (home) bias of European central bankers: new evidence based on speeches. *Applied Economics*, 49:1–18, August 2016. doi: 10.1080/00036846.2016.1210782.

Hamza Bennani, Nicolas Fanta, Pavel Gertler, and Roman Horvath. Does central bank communication signal future monetary policy in a (post)-crisis era? The case of the ECB. *Journal of International Money and Finance*, 104:102167, June 2020. ISSN 0261-5606. doi: 10.1016/j.jimonfin.2020.102167. URL https://www.sciencedirect.com/science/article/pii/S0261560618302882.

Kerstin Bernoth and Geraldine Dany-Knedlik. The ECB's Communication Strategy: Limits and Challenges After the Financial Crisis. January 2020. URL https://policycommons.net/artifacts/1337418/the-ecbs-communication-strategy/1945266/. Publisher: <bound method Organization.get_name_with_acronym of <Organization: European Parliamentary Research Service>>.

Carola Binder. Fed speak on main street: Central bank communication and household expectations. *Journal of Macroeconomics*, 52:238–251, June 2017. ISSN 0164-0704. doi: 10.1016/j.jmacro.2017.05.003. URL https://www.sciencedirect.com/science/article/pii/S0164070416302312.

André Binette and Dmitri Tchebotarev. Canada's Monetary Policy Report: If Text Could Speak, What Would It Say? 2019.

Alan S. Blinder, Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen. Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence. *Journal of Economic Literature*, 46(4):910–945, December 2008. ISSN 0022-0515. doi: 10.1257/jel.46.4.910. URL https://www.aeaweb.org/articles?id=10.1257/jel.46.4.910.

Alan S Blinder, Michael Ehrmann, Jakob de Haan, and David-Jan Jansen. Central Bank Communication with the General Public: Promise or False Hope? 2022.

Martin T. Bohl, Dimitrios Kanelis, and Pierre L. Siklos. Central bank mandates: How differences can influence the content and tone of central bank communication. *Journal of International Money and Finance*, 130:102752, 2023. ISSN 0261-5606. doi: https://doi.org/10.

1016/j.jimonfin.2022.102752. URL https://www.sciencedirect.com/science/article/pii/S0261560622001553.

Michael Bordo and Klodiana Istrefi. Perceived FOMC: The making of hawks, doves and swingers. *Journal of Monetary Economics*, 136:125–143, May 2023. ISSN 03043932. doi: 10.1016/j.jmoneco.2023.03.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0304393223000272.

Benjamin Born, Michael Ehrmann, and Marcel Fratzscher. Central Bank Communication on Financial Stability. *The Economic Journal*, 124(577):701–734, June 2014. ISSN 0013-0133. doi: 10.1111/ecoj.12039. URL https://doi.org/10.1111/ecoj.12039.

Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior Research Methods*, 46(3):904—-911, 2013.

Emanuele Campiglio, Jérôme Deyris, Davide Romelli, and Ginevra Scalisi. Warning words in a warming world: Central bank communication and climate change. *European Economic Review*, 178:105101, September 2025. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2025.105101. URL https://www.sciencedirect.com/science/article/pii/S0014292125001515.

Sarah E. Cannon. Committee decisions, voting records, and future policy. *Journal of Money, Credit and Banking*, 47(5):901–926, 2015.

Jan Celler. Readability and sentiment analysis of central bank communication in central and eastern europe. 28(4):1018–1033, 2024. doi: 10.20965/jaciii.2024.p1018.

Carl Andreas Claussen, Øistein Røisland, and Tommy Sveen. Words matter: The role of communication in monetary policy. *Journal of Money, Credit and Banking*, 44(4):719–744, 2012.

Michael P. Clements and J. James Reade. Forecasting and forecast narratives: The Bank of England Inflation Reports. *International Journal of Forecasting*, 36(4):1488–1500, October 2020. ISSN 0169-2070. doi: 10.1016/j.ijforecast.2019.08.013. URL https://www.sciencedirect.com/science/article/pii/S0169207019302377.

Olivier Coibion and Yuriy Gorodnichenko. Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation. *American Economic Journal: Macroeconomics*, 7

(1):197–232, January 2015. ISSN 1945-7707. doi: 10.1257/mac.20130306. URL https://www.aeaweb.org/articles?id=10.1257/mac.20130306.

Olivier Coibion, Dimitris Georgarakos, Yuriy Gorodnichenko, and Maarten van Rooij. How Does Consumption Respond to News about Inflation? Field Evidence from a Randomized Control Trial. *American Economic Journal: Macroeconomics*, 15(3):109–152, July 2023. ISSN 1945-7707. doi: 10.1257/mac.20200445. URL https://www.aeaweb.org/articles?id=10.1257/mac.20200445.

Max Coltheart. The mrc psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, 33(4):497–505, 1981. doi: 10.1080/14640748108400805. URL https://doi.org/10.1080/14640748108400805.

Philippine Cour-Thimann and Alexander Jung. Interest-rate setting and communication at the ECB in its first twenty years. *European Journal of Political Economy*, 70:102039, December 2021. ISSN 0176-2680. doi: 10.1016/j.ejpoleco.2021.102039. URL https://www.sciencedirect.com/science/article/pii/S0176268021000409.

Carin van der Cruijsen and Anna Samarina. Trust in the ECB in turbulent times. *Working Papers*, July 2021. URL https://ideas.repec.org//p/dnb/dnbwpp/722.html. Number: 722 Publisher: DNB.

Alex Cukierman and Allan Meltzer. A theory of ambiguity, credibility, and inflation under discretion and asymmetric information. *Econometrica*, 54:1099–1128, 02 1986. doi: 10.2307/1912324.

Jakob de Haan and Lex Hoogduin. ECB communication policies: An overview and comparison with the Federal Reserve. *Journal of International Money and Finance*, 142:103050, April 2024. ISSN 0261-5606. doi: 10.1016/j.jimonfin.2024.103050. URL https://www.sciencedirect.com/science/article/pii/S0261560624000378.

Milena Djourelova, Michael McMahon, and Ryland Thomas. Fomc communication and the heterogeneity of monetary policy preferences. *Journal of Monetary Economics*, 2025. Forthcoming.

Thang Ngoc Doan, Dong Phu Do, and Dat Van Luong. Monetary stance and favorableness of the monetary policy in the media: the case of Vietnam. *Journal of Asian Business and Economic*

Studies, 31(2):111–123, January 2023. ISSN 2515-964X. doi: 10.1108/JABES-02-2023-0038. URL https://doi.org/10.1108/JABES-02-2023-0038. Publisher: Emerald Publishing Limited.

Asad Dossani. Central bank tone and currency risk premia. *Journal of International Money and Finance*, 117:102424, October 2021. ISSN 0261-5606. doi: 10.1016/j.jimonfin.2021.102424. URL https://www.sciencedirect.com/science/article/pii/S0261560621000759.

Lena Dräger, Michael J. Lamla, and Damjan Pfajfar. Are survey expectations theory-consistent? The role of central bank communication and news. *European Economic Review*, 85:84–111, June 2016. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2016.01.010. URL https://www.sciencedirect.com/science/article/pii/S0014292116300149.

Ioana Duca-Radu, Geoff Kenny, and Andreas Reuter. Inflation expectations, consumption and the lower bound: Micro evidence from a large multi-country survey. *Journal of Monetary Economics*, 118(C):120–134, 2021. URL https://ideas.repec.org//a/eee/moneco/v118y2021icp120-134.html. Publisher: Elsevier.

Michael Ehrmann and Marcel Fratzscher. Explaining Monetary Policy in Press Conferences. *International Journal of Central Banking*, 5:42–84, June 2009. doi: 10.2139/ssrn.989997.

Michael Ehrmann and Jonathan Talmi. Starting from a blank page? Semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, 111:48–62, May 2020. ISSN 0304-3932. doi: 10.1016/j.jmoneco.2019.01.028. URL https://www.sciencedirect.com/science/article/pii/S0304393219300169.

Albert Einstein. Zur elektrodynamik bewegter körper. *Annalen der Physik*, 322:891–921, 1905.

Patrick Eugster and Matthias W. Uhl. Forecasting inflation using sentiment. *Economics Letters*, 236:111575, March 2024. ISSN 0165-1765. doi: 10.1016/j.econlet.2024.111575. URL https://www.sciencedirect.com/science/article/pii/S0165176524000582.

Federal Reserve Bank of St. Louis. Fomc speak archive. https://fraser.stlouisfed.org/timeline/fomc-speak-archive, 2024. Accessed: YYYY-MM-DD.

Kristin Forbes, Jongrim Ha, and M. Ayhan Kose. Tradeoffs over Rate Cycles: Activity, Inflation and the Price Level. Technical Report w33825, National Bureau of Economic Research, May 2025. URL https://www.nber.org/papers/w33825.

Pavel Gertler and Roman Horvath. Central bank communication and financial markets: New high-frequency evidence. *Journal of Financial Stability*, 36:336–345, June 2018. ISSN 1572-3089. doi: 10.1016/j.jfs.2018.03.002. URL https://www.sciencedirect.com/science/article/pii/S1572308918301645.

Daniel Gros and Angela Capolongo. Reaching a Wider Audience: Is the ECB Trending? 2020.

Chen Gu, Denghui Chen, Raluca Stan, and Aizhong Shen. It is not just *What* you say, but *How* you say it: Why tonality matters in central bank communication. *Journal of Empirical Finance*, 68:216–231, September 2022. ISSN 0927-5398. doi: 10.1016/j.jempfin.2022.07.008. URL https://www.sciencedirect.com/science/article/pii/S0927539822000561.

Anne Lundgaard Hansen and Sophia Kazinnik. Can ChatGPT Decipher Fedspeak?, April 2024. URL https://papers.ssrn.com/abstract=4399406.

Stephen Hansen, Michael McMahon, and Matthew Tong. The long-run information effect of central bank communication. *Journal of Monetary Economics*, 108:185–202, December 2019. ISSN 0304-3932. doi: 10.1016/j.jmoneco.2019.09.002. URL https://www.sciencedirect.com/science/article/pii/S0304393219301606.

Jochen Hartmann. Emotion english distilroberta-base. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/, 2022.

Bernd Hayo and Johannes Zahner. What is that noise? Analysing sentiment-based variation in central bank communication. *Economics Letters*, 222:110962, January 2023. ISSN 0165-1765. doi: 10.1016/j.econlet.2022.110962. URL https://www.sciencedirect.com/science/article/pii/S0165176522004360.

Bernd Hayo, Kai Henseler, Marc Steffen Rapp, and Johannes Zahner. Complexity of ECB communication and financial market trading. *Journal of International Money and Finance*,

128:102709, November 2022. ISSN 0261-5606. doi: 10.1016/j.jimonfin.2022.102709. URL https://www.sciencedirect.com/science/article/pii/S0261560622001127.

David-Jan Jansen and Jakob de Haan. An Assessment of the Consistency of ECB Communication. May 2013. doi: 10.7551/mitpress/9470.003.0014. URL https://direct.mit.edu/books/edited-volume/3709/chapter/123202/An-Assessment-of-the-Consistency-of-ECB.

Kohei Kawamura, Yohei Kobashi, Masato Shizume, and Kozo Ueda. Strategic central bank communication: Discourse analysis of the Bank of Japan's Monthly Report. *Journal of Economic Dynamics and Control*, 100:230–250, March 2019. ISSN 0165-1889. doi: 10.1016/j.jedc.2018.11.007. URL https://www.sciencedirect.com/science/article/pii/S0165188919300065.

K.L. Kliesen, B. Levine, and C.J. Waller. Gauging market responses to monetary policy communication. *Federal Reserve Bank of St. Louis Review*, 101(2):69–91, 2019. doi: 10.20955/r.101.69-91.

Pongsak Luangaram and Warapong Wongwachara. More Than Words: A Textual Analysis of Monetary Policy Communication. 2017.

Tiff Macklem and Jill Vardy. 20 years of central bank communications, and lessons for the future. In *Central banking in the Americas: Lessons from two decades*, volume 127 of *BIS Papers*, pages 55–67. Bank for International Settlements, 2023.

Ulrike Malmendier, Stefan Nagel, and Zhen Yan. The making of hawks and doves. *Journal of Monetary Economics*, 117:19–42, 2021. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2020.04.002. URL https://www.sciencedirect.com/science/article/pii/S0304393220300490.

Elizabeth E. Meade and Nathan Sheets. Regional influences on FOMC voting patterns. *Journal of Money, Credit and Banking*, 37(4):661–677, 2005.

Apel Mikael and Grimaldi Marianna Blix. How Informative Are Central Bank Minutes? *Review of Economics*, 65(1):53–76, 2014. URL https://ideas.repec.org//a/lus/reveco/v65y2014i1p53-76.html. Publisher: De Gruyter.

Andy Moniz and Franciska de Jong. Predicting the Impact of Central Bank Communications on Financial Market Investors' Interest Rate Expectations. In Valentina Presutti, Eva Blomqvist,

Raphael Troncy, Harald Sack, Ioannis Papadakis, and Anna Tordai, editors, *The Semantic Web: ESWC 2014 Satellite Events*, pages 144–155, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11955-7. doi: 10.1007/978-3-319-11955-7_12.

Conor Parle. The financial market impact of ECB monetary policy press conferences — A text based approach. *European Journal of Political Economy*, 74:102230, September 2022. ISSN 0176-2680. doi: 10.1016/j.ejpoleco.2022.102230. URL https://www.sciencedirect.com/science/article/pii/S0176268022000428.

Alberto Parola, Jessica Mary Lin, Arndis Simonsen, Vibeke Bliksted, Yuan Zhou, Huiling Wang, Lana Inoue, Katja Koelkebeck, and Riccardo Fusaroli. Speech disturbances in schizophrenia: Assessing cross-linguistic generalizability of NLP automated measures of coherence. 259:59–70, 2023. ISSN 0920-9964. doi: 10.1016/j.schres.2022.07.002. URL https://www.sciencedirect.com/science/article/pii/S0920996422002742.

Moritz Pfeifer and Vincent P. Marohl. CentralBankRoBERTa: A fine-tuned large language model for central bank communications. *The Journal of Finance and Data Science*, 9:100114, November 2023. ISSN 2405-9188. doi: 10.1016/j.jfds.2023.100114. URL https://www.sciencedirect.com/science/article/pii/S2405918823000302.

Matthieu Picault and Thomas Renault. Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance*, 79:136–156, December 2017. ISSN 0261-5606. doi: 10.1016/j.jimonfin.2017.09.005. URL https://www.sciencedirect.com/science/article/pii/S0261560617301808.

Jack W. Rae, Sebastian Borgeaud, Trevor Cai, Katie Millican, Jordan Hoffmann, Francis Song, John Aslanides, Sarah Henderson, Roman Ring, Susannah Young, Eliza Rutherford, Tom Hennigan, Jacob Menick, Albin Cassirer, Richard Powell, George van den Driessche, Lisa Anne Hendricks, Maribeth Rauh, Po-Sen Huang, Amelia Glaese, Johannes Welbl, Sumanth Dathathri, Saffron Huang, Jonathan Uesato, John Mellor, Irina Higgins, Antonia Creswell, Nat McAleese, Amy Wu, Erich Elsen, Siddhant Jayakumar, Elena Buchatskaya, David Budden, Esme Sutherland, Karen Simonyan, Michela Paganini, Laurent Sifre, Lena Martens, Xiang Lorraine Li, Adhiguna Kuncoro, Aida Nematzadeh, Elena Gribovskaya, Domenic Donato, Angeliki Lazaridou, Arthur

Mensch, Jean-Baptiste Lespiau, Maria Tsimpoukelli, Nikolai Grigorev, Doug Fritz, Thibault Sottiaux, Mantas Pajarskas, Toby Pohlen, Zhitao Gong, Daniel Toyama, Cyprien de Masson d'Autume, Yujia Li, Tayfun Terzi, Vladimir Mikulik, Igor Babuschkin, Aidan Clark, Diego de Las Casas, Aurelia Guy, Chris Jones, James Bradbury, Matthew Johnson, Blake Hechtman, Laura Weidinger, Iason Gabriel, William Isaac, Ed Lockhart, Simon Osindero, Laura Rimell, Chris Dyer, Oriol Vinyals, Kareem Ayoub, Jeff Stanway, Lorrayne Bennett, Demis Hassabis, Koray Kavukcuoglu, and Geoffrey Irving. Scaling language models: Methods, analysis & insights from training gopher, 2022. URL http://arxiv.org/abs/2112.11446.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer, 2023. URL http://arxiv.org/abs/1910.10683.

Alessandro Riboni and Francisco Ruge-Murcia. Membership turnover and policy disagreement at the fomc. Working Papers hal-05229751, HAL, Aug 2025. URL https://ideas.repec.org/p/hal/wpaper/hal-05229751.html.

Maik Schmeling and Christian Wagner. Does Central Bank Tone Move Asset Prices?, July 2023. URL https://papers.ssrn.com/abstract=2629978.

Graham G. Scott, Anne Keitel, Marc Becirspahic, Bo Yao, and Sara C. Sereno. The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3):1258–1270, June 2019. ISSN 1554-3528. doi: 10.3758/s13428-018-1099-3. URL https://doi.org/10.3758/s13428-018-1099-3.

Agam Shah, Suvan Paturi, and Sudheer Chava. Trillion dollar words: A new financial dataset, task & market analysis. *Available at SSRN 4447632*, 2023.

Thiago Christiano Silva, Kei Moriya, and Romain M. Veyrune. From Text to Quantified Insights – A Large-Scale LLM Analysis of Central Bank Communication. *IMF Working Papers*, 2025 (109), June 2025. doi: 10.5089/9798229013802.001.A001. URL https://www.elibrary.imf.org/view/journals/001/2025/109/article-A001-en.xml. ISBN: 9798229013802.

L. A. Smales and N. Apergis. Does more complex language in FOMC decisions impact fi-

nancial markets? *Journal of International Financial Markets, Institutions and Money*, 51: 171–189, November 2017. ISSN 1042-4431. doi: 10.1016/j.intfin.2017.08.003. URL https://www.sciencedirect.com/science/article/pii/S104244311630138X.

Peter T. Hughes and Stefan Kesting. A literature review on central bank communication. *On the Horizon*, 22(4):328–340, January 2014. ISSN 1074-8121. doi: 10.1108/OTH-07-2014-0027. URL https://doi.org/10.1108/OTH-07-2014-0027. Publisher: Emerald Group Publishing Limited.

Peter Tillmann. What do FOMC dissenters tell us about monetary policy? *Journal of Macroeconomics*, 33(4):684–699, 2011.

Peter Tillmann and Andreas Walter. The effect of diverging communication: The case of the ECB and the Bundesbank. *Economics Letters*, 176:68–74, March 2019. ISSN 0165-1765. doi: 10.1016/j.econlet.2018.12.035. URL https://www.sciencedirect.com/science/article/pii/S0165176518305299.

C. van der Cruijsen, D.-J. Jansen, and J. de Haan. How much does the public know about the ECB's monetary policy? Evidence from a survey of dutch households. *International Journal of Central Banking*, 11(4):169–218, 2015.

Renske Van der Cruijsen, Sabine Peters, Kelly P. M. Zoetendaal, Jennifer H. Pfeifer, and Eveline A. Crone. Direct and reflected self-concept show increasing similarity across adolescence: A functional neuroimaging study. *Neuropsychologia*, 129:407–417, June 2019. ISSN 0028-3932. doi: 10.1016/j.neuropsychologia.2019.05.001. URL https://www.sciencedirect.com/science/article/pii/S0028393219301101.

7 Appendix

7.1 Scatter Graphs: Board Members versus Bank Presidents

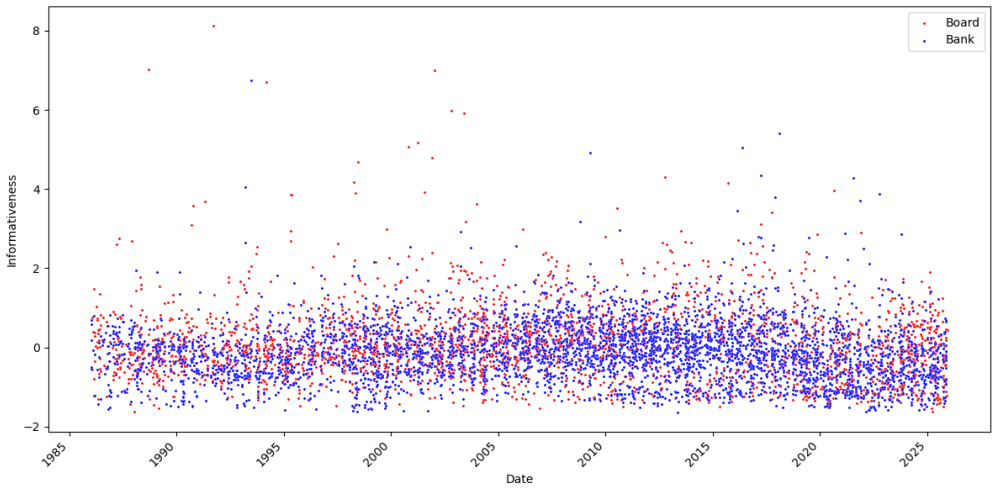

Figure 12: Board Members versus Bank Presidents: Informativeness

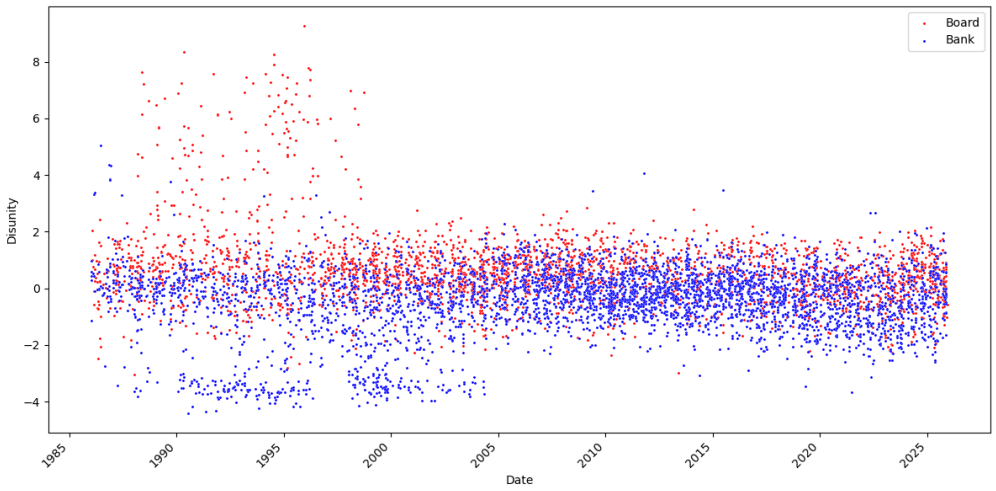

Figure 13: Board Members versus Bank Presidents: Disunity

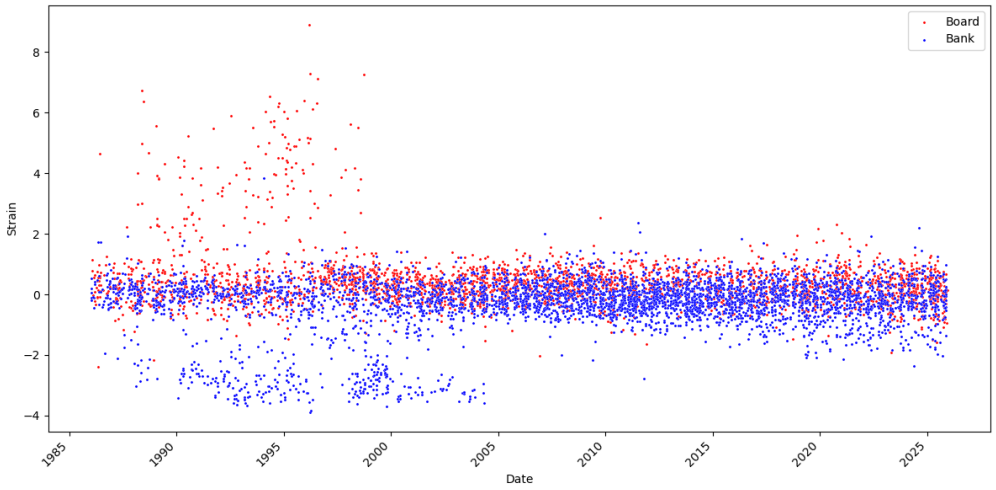

Figure 14: Board Members versus Bank Presidents: Cognitive Strain

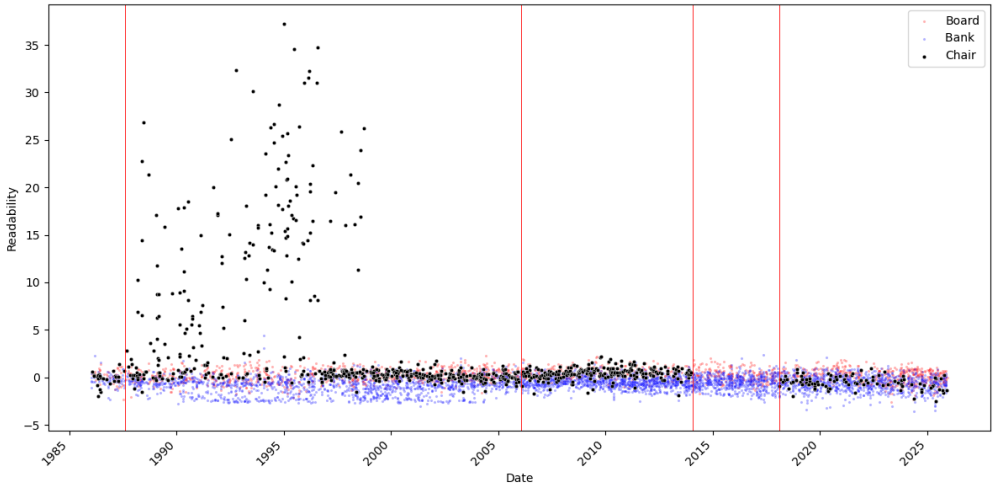

Figure 15: Board Members - Chair: Readability

7.2 Scatter Graphs: Within Board Members

7.3 Details on the 12 District Bank Presidents

7.3.1 Federal Reserve Bank of New York

Unlike most Reserve Banks, the New York Fed publishes speeches not only from its president but also from several senior officials. This practice reflects its unique institutional role, including responsibility for the System Open Market Account (SOMA), daily market operations, and a large research and markets group that engages in frequent public communication.

Speech frequency is relatively stable across the sample, with a noticeable increase after 2009,

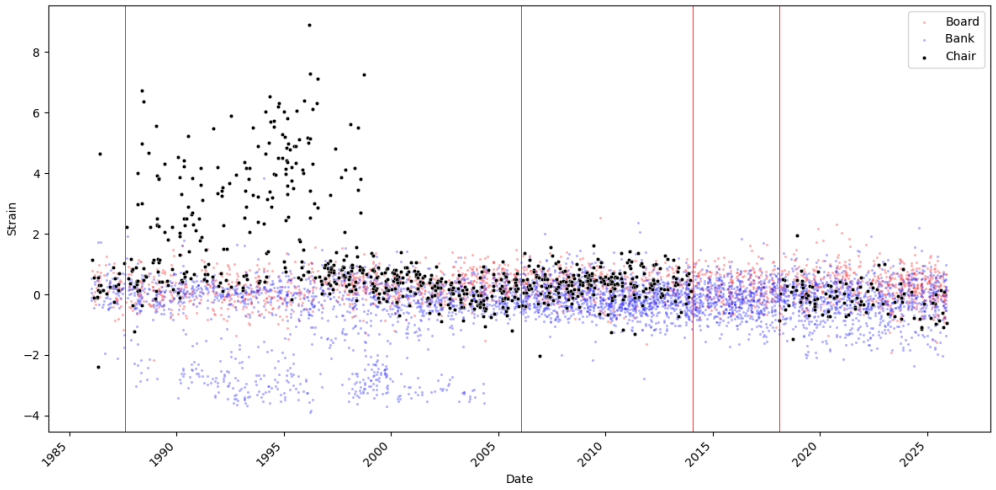

Figure 16: Board Members - Chair: Cognitive Strain

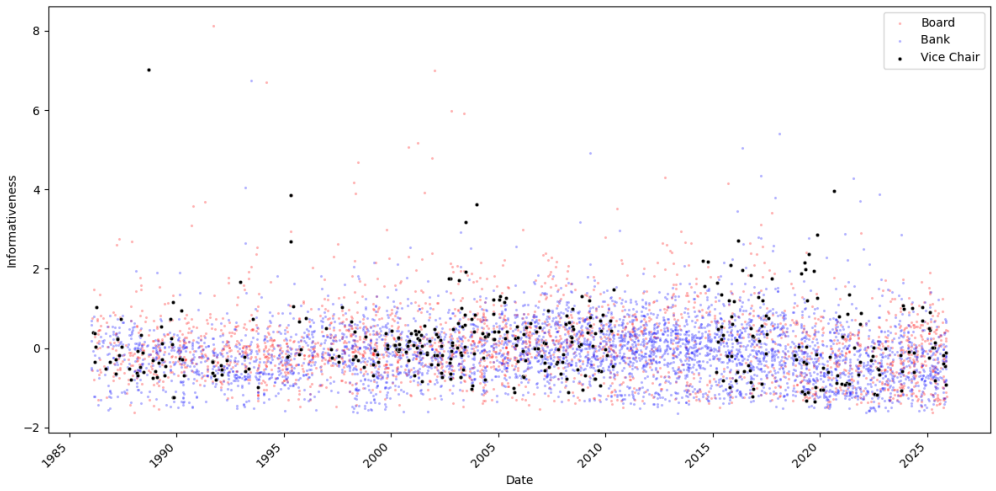

Figure 17: Board Members - Vice Chair: Informativeness

consistent with the broader post-GFC rise in public communication across the Federal Reserve System. Abstractness declines after 2020, which aligns with the shift toward more practical, operational communication during the pandemic period (see Figure 19).

In general, the presidents' speeches are fairly uniform in style, with the exception of John Williams. His tenure overlaps with the post-pandemic period, during which communication across the System became more pragmatic and accessible, which may explain the slight deviation despite his academic background. Timothy Geithner exhibits slightly higher (though still modest) abstractness and disunity, which is consistent with his diverse professional background and the unusually turbulent period during which he served.

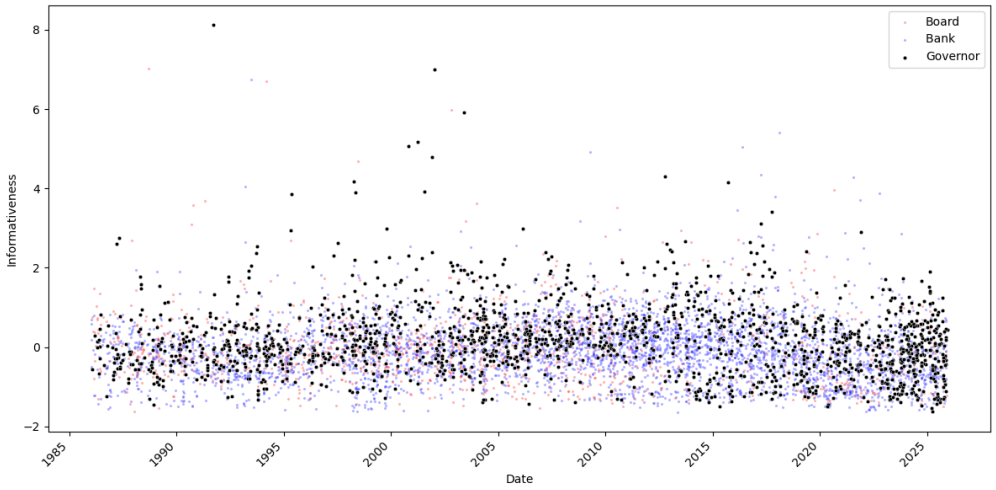

Figure 18: Board Members - other Governors: Informativeness

Name	**Tenure**	**Education (Field + Institution)**	**Relevant Professional Background**
E. Gerald Corrigan	1985–1993	BA in Economics, Fairfield University; MA & PhD in Economics, Fordham University	Career economist within the Federal Reserve System; strong experience dealing with financial and bank crisis.
William J. McDonough	1993–2003	BA in Economics, College of the Holy Cross	Long career in international banking (First Chicago Corp.).
Timothy F. Geithner	2003–2009	BA in Government/Asian Studies, Dartmouth College; MA in International Economics & East Asian Studies, Johns Hopkins SAIS	Senior Treasury official; IMF policy director; expertise in international finance and crisis response; later U.S. Treasury Secretary.
William C. Dudley	2009–2018	BA in Economics, New College of Florida; PhD in Economics, University of California, Berkeley	Chief U.S. Economist at Goldman Sachs; partner and managing director.
John C. Williams	2018–present	BA in Economics, UC Berkeley; MSc in Economics, London School of Economics; PhD in Economics, Stanford University	Academic economist; research on monetary policy rules; former President of the San Francisco Fed; expertise in macroeconomic modeling.

Table 9: Presidents of the Federal Reserve Bank of New York (1985–2025): Education and Professional Background

One unusual feature appears toward the end of William Dudley's tenure (2009–2018): a small cluster of speeches with unusually high informativeness, without corresponding increases in other complexity metrics. This pattern likely reflects multi-topic, information-dense speeches rather than a broader shift in communication style. Given the New York Fed's unique mandates—such

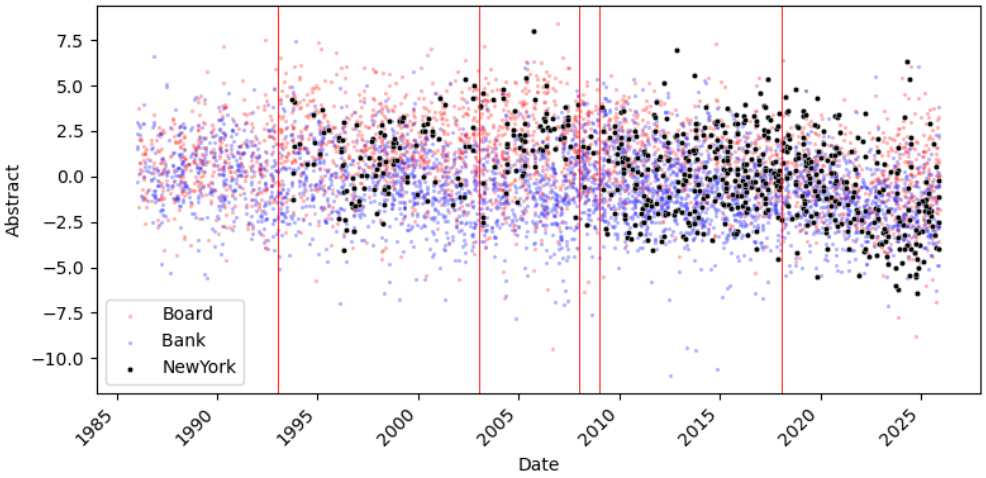

Figure 19: Federal Reserve Bank of New York Abstractness

as implementing quantitative easing, managing the SOMA portfolio, and overseeing balance sheet normalization—it is plausible that certain speeches during this period covered a wider range of technical issues than usual.

Because the New York Fed is institutionally heterogeneous, I also re-estimated the results using only speeches by Bank Presidents. The main differences arise from the expected reduction in total speech count, and the patterns observed for Timothy Geithner become slightly more pronounced. Otherwise, the overall results remain consistent.

7.3.2 Federal Reserve Bank of Boston

Readability, abstractness, and cognitive strain (especially in the upper-bound measure) are slightly lower prior to 2007, which corresponds to Cathy Minehan's tenure. One possible explanation is that Minehan's background in political science and finance—rather than economics—may have contributed to a somewhat more accessible communication style, although this interpretation should be viewed cautiously. Other complexity measures show no strong deviations across presidents. The increase in complexity after 2007 may also reflect the broader communication environment shaped by the Great Financial Crisis (GFC), which required more detailed and analytically dense public explanations from Federal Reserve officials.

A pattern consistent with what is observed in New York also appears here: during the post-pandemic period, presidential speeches become more accessible in terms of complexity. In Boston's case, this

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Frank E. Morris	1968–1988	BSc in Engineering, General Motors Institute; BA in Economics, Wayne University	Economist at the Office of Price Stabilization and CIA; senior Fed economist; long-serving Boston Fed president with strong macroeconomic and policy background.
Richard F. Syron	1989–1994	BA in Economics, Boston College; MA & PhD in Economics, Tufts University	An economist, officer, and economic adviser in the Research Department; active in public policy debates.
Cathy E. Minehan	1994–2007	BA in Political Science, University of Rochester; MBA (Finance), New York University	Career Federal Reserve official; extensive experience in payments systems and bank supervision.
Eric S. Rosengren	2007–2021	BA in Economics, Colby College; PhD in Economics, University of Wisconsin–Madison	Research economist at Boston Fed; specialist in macroeconomics, international banking, bank supervision, and risk management.
Susan M. Collins	2022–present	BA in Economics, Harvard University; PhD in Economics, MIT	Academic economist; provost and dean at University of Michigan; research in international macroeconomics; former director at Chicago Fed.

Table 10: Presidents of the Federal Reserve Bank of Boston (1986–2025): Education and Professional Background

occurs under Susan Collins, despite her academic background, aligning with the System-wide shift toward clearer, more practical communication during and after the pandemic. The figures for abstractness and the upper bound of cognitive strain are shown in Figures 20 and 21.

7.3.3 Federal Reserve Bank of Philadelphia

Speech frequency is relatively consistent over time, with the archive becoming reliably populated only after the late 1990s. Across the complexity measures, abstractness is notably higher during Charles Plosser's presidency (2006–2015), with Anthony Santomero (2000–2006) falling in an intermediate position and Patrick Harker (2015–present) exhibiting the lowest levels. The same ordering appears in disunity, informativeness (Shannon entropy), and readability difficulty. Cognitive strain does not follow this pattern, although its upper-bound version shows a mild resemblance. Figures 22 and 23 illustrate the patterns for abstractness and disunity

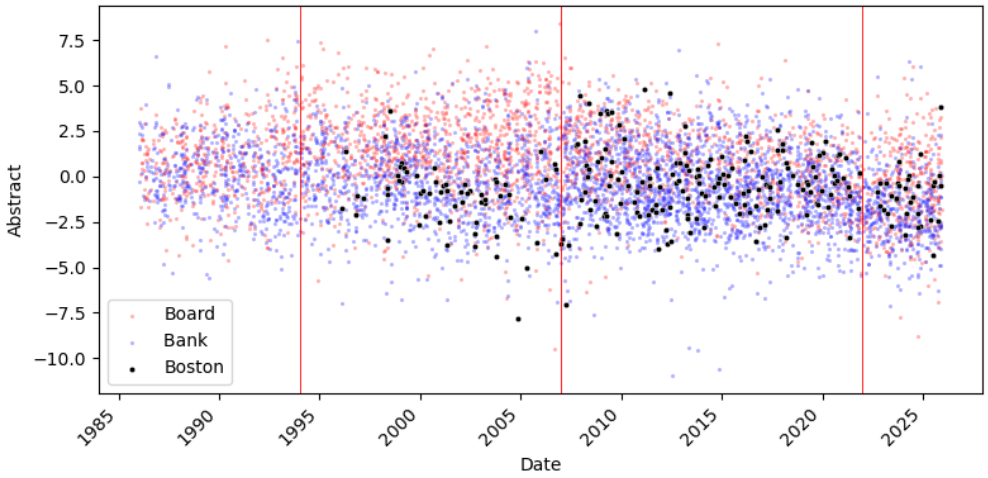

Figure 20: Federal Reserve Bank of Boston Abstractness

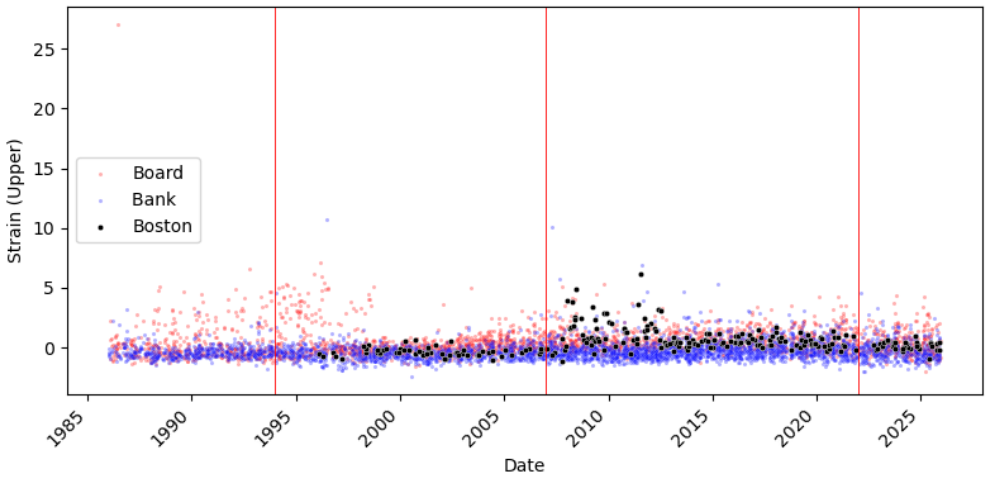

Figure 21: Federal Reserve Bank of Boston Cognitive Strain (Upper Bound)

Plosser's communication style aligns well with these results. As an academic macroeconomist trained at the University of Rochester, his speeches frequently addressed technical topics such as monetary policy frameworks, long-run inflation control, credibility and commitment, and structural macroeconomic issues. These themes naturally lend themselves to higher abstractness, broader vocabulary, and more complex structure than those of the other two presidents.

Santomero, by contrast, was a financial economist whose work focused more on banking, risk, and financial institutions. His communication style is therefore less theoretical than Plosser's but still more technical than Harker's, which is consistent with his intermediate position in the complexity metrics.

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Edward G. Boehne	1981–2000	BA in Economics, Indiana University; PhD in Economics, University of Pennsylvania	Career Federal Reserve economist; specialist in regional economics; long-serving policymaker with deep experience in monetary policy and economic research.
Anthony M. Santomero	2000–2006	BS in Economics, Fordham University; MA & PhD in Economics, Brown University	Academic economist; professor at Wharton; expert in banking, financial institutions, and risk management.
Charles I. Plosser	2006–2015	BS in Engineering, Vanderbilt University; PhD in Economics, University of Chicago	Academic macroeconomist; dean of the Simon School (University of Rochester); research in real business cycles and monetary policy rules.
Patrick T. Harker	2015–2025	BSE, MSE, PhD in Engineering, University of Pennsylvania; MA in Economics, University of Pennsylvania	University president (University of Delaware); operations researcher; academic administrator.

Table 11: Presidents of the Federal Reserve Bank of Philadelphia (1986–2025): Education and Professional Background

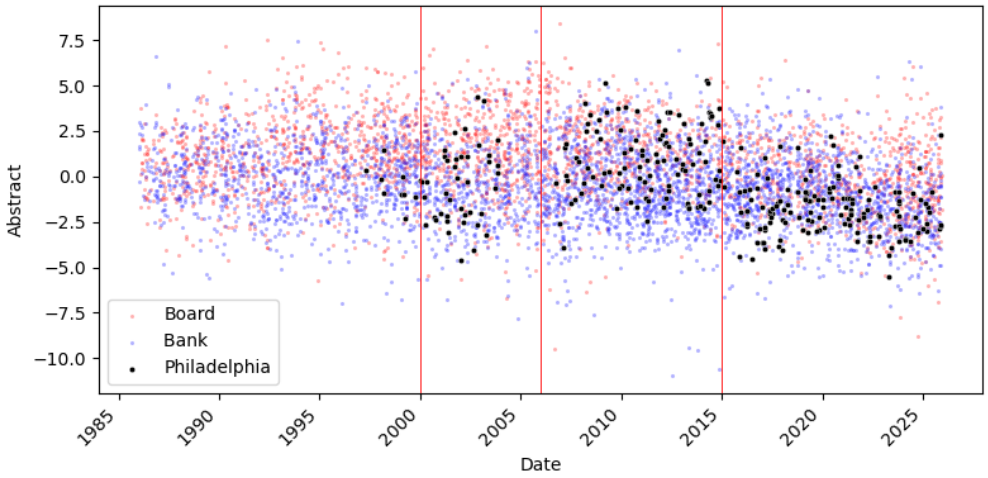

Figure 22: Federal Reserve Bank of Philadelphia Abstractness

Harker's background includes training in engineering and operations research, as well as serving as president of the University of Delaware. This combination tends to produce a communication style that is more concrete and managerial in tone, with less emphasis on theoretical macroeconomic framing. This aligns with the lower abstractness, lower disunity, and easier readability observed during his tenure.

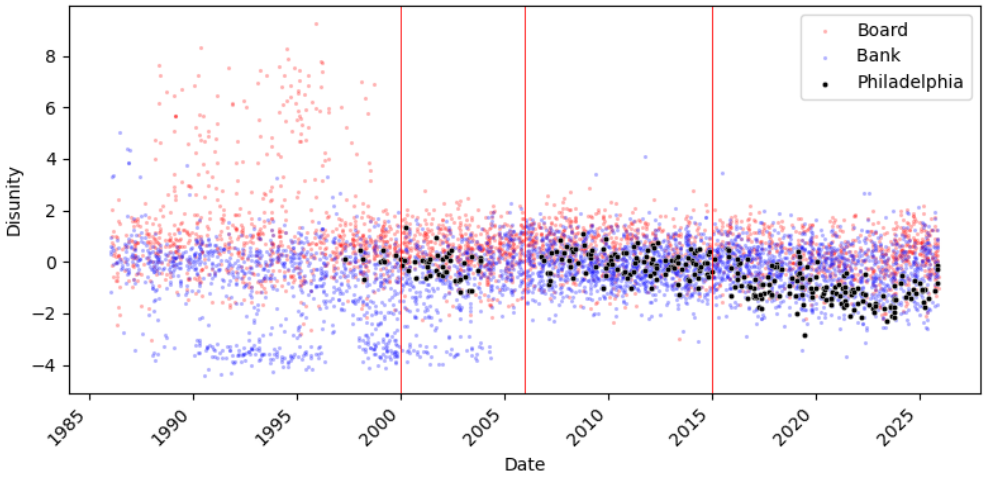

Figure 23: Federal Reserve Bank of Philadelphia Disunity

7.3.4 Federal Reserve Bank of Cleveland

The number of speeches appears broadly consistent over time, with the exception of a gradual decline between roughly 2000 and 2005. One possible explanation is that, during this period of increasing Federal Reserve transparency, there may have been less need for the same volume of public speeches as in the early years of Jerry Jordan's tenure, when he may have been more focused on establishing his policy stance. This could help account for the reduced speech frequency toward the end of his term, although this interpretation should be viewed cautiously.

Abstractness (Figure 24) is somewhat difficult to interpret directly from the graph, but several patterns appear plausible. Jordan's speeches (1992–2003) seem to exhibit higher abstractness than those of other presidents. This aligns with his background in macroeconomic theory and monetary policy, which often involves more conceptual or theoretical themes. In contrast, Sandra Pianalto's communication may have been more pragmatic and operational in tone, consistent with her training in management and business. Abstractness then rises again under Loretta Mester (2014–2024), which is consistent with her research-oriented background and academic training, before declining noticeably during the pandemic period. This decline likely reflects the broader System-wide shift toward more practical, operational communication during the crisis and the early tenure of Beth Hammack.

For the other complexity measures—disunity, informativeness, readability difficulty, and cognitive strain—there is a general upward trend in complexity from approximately 2003 to 2024,

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Lee Hoskins	1987–1991	BA in Economics, University of California, Berkeley; PhD in Economics, University of California, Los Angeles	Economist at the Federal Reserve Bank of Minneapolis; research in monetary policy and financial regulation; known for strong anti-inflation stance.
Jerry L. Jordan	1992–2003	BA in Economics, California State University, Northridge; PhD in Economics, University of California, Los Angeles	Senior economist at the Federal Reserve Bank of St. Louis; private-sector economist; expertise in monetary policy, banking, and macroeconomic theory.
Sandra Pianalto	2003–2014	BA in Economics, University of Akron; MA in Economics, George Washington University; Advanced Management Program at Duke University's Fuqua School of Business	Career Federal Reserve economist; staff of the Budget Committee of the US House of Representatives.
Loretta J. Mester	2014–2024	BA in Mathematics & Economics, Barnard College (Columbia University); MA & PhD in Economics, Princeton University	Executive VP and Director of Research at the Philadelphia Fed; academic economist; specialist in banking, financial stability, and monetary policy.
Beth M. Hammack	2024–present	BA in Quantitative economics & History, Stanford University	Former Goldman Sachs partner; global finance executive; extensive experience in capital markets, risk management, and Treasury financing.

Table 12: Presidents of the Federal Reserve Bank of Cleveland (1986–2025): Education and Professional Background

Figure 24: Federal Reserve Bank of Cleveland Abstractness

although earlier periods show greater volatility. Informativeness may decline slightly after the pandemic, but the pattern is not definitive. These trends can be partly interpreted through the same contextual factors discussed above. However, unlike abstractness, these measures do not show a pronounced decline associated with economic conditions; instead, they remain relatively stable (and in some cases elevated) even during the pandemic period.

7.3.5 Federal Reserve Bank of Richmond

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Robert P. Black	1973–1992	BA, MA, PhD in Economics, University of Virginia	Career Federal Reserve economist; specialist in monetary policy and bank research; long tenure shaped by strong anti-inflation views.
J. Alfred Broaddus Jr.	1993–2004	MA & PhD in Economics, University of Virginia	Director of Research at the Richmond Fed; over two decades in Fed research; expertise in macroeconomics and monetary policy formulation.
Jeffrey M. Lacker	2004–2017	PhD in Economics, University of Wisconsin–Madison	Academic economist (Purdue University faculty); Director of Research at Richmond Fed; specialist in financial stability, payments systems, and macroeconomic theory.
Thomas I. Barkin	2018–present	BA in Economics; JD/MBA, Harvard University	Senior partner at McKinsey & Company; former CFO and chief risk officer; extensive experience in management, finance, and organizational strategy.

Table 13: Presidents of the Federal Reserve Bank of Richmond (1986–2025): Education and Professional Background

Speech frequency gradually increases over time, with a noticeable rise beginning under Jeffrey Lacker (2004–2018). The archive becomes reliably populated only after the mid-1990s, which explains the limited number of speeches available for J. Alfred Broaddus (1993–2004). His smaller sample makes it difficult to draw firm conclusions, though his complexity measures appear to hover around the district average. For the two presidents since 2004, the scatterplots for disunity and cognitive strain are shown in Figures 25 and 26.

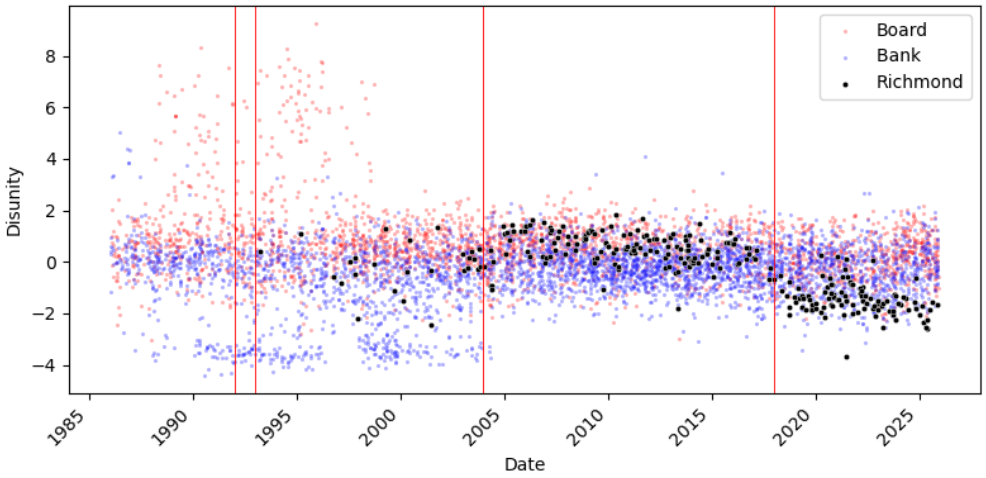

Figure 25: Federal Reserve Bank of Richmond Disunity

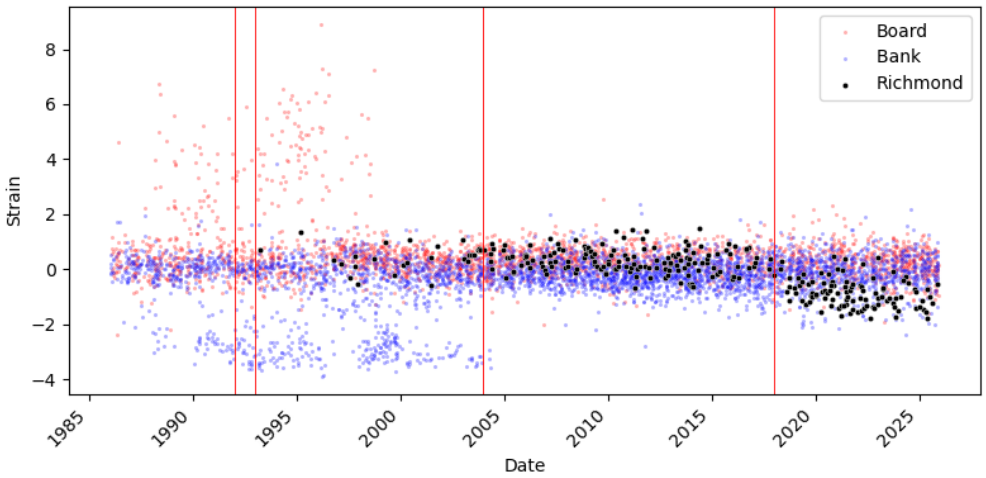

Figure 26: Federal Reserve Bank of Richmond Coginitive Strain

Across all complexity dimensions—abstractness, disunity, informativeness, readability difficulty, and cognitive strain—Lacker's speeches stand out as substantially more complex than those of his successor, Thomas Barkin (2018–present). This pattern is consistent with Lacker's background. He is widely regarded as one of the most academically oriented and theoretically driven Reserve Bank presidents of the modern era, and his speeches frequently incorporated macroeconomic theory and analytical discussions. These themes naturally generate higher abstractness, broader vocabulary, and more intricate structure. Barkin's communication style contrasts sharply with this. Coming from McKinsey, he brings a business and management background and tends to adopt a practical, operational approach to communication. His speeches are more concrete, more focused

on regional and organizational issues, and generally easier to read. This aligns with the lower complexity observed across all measures during his tenure.

7.3.6 Federal Reserve Bank of Chicago

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Silas Keehn	1981–1994	BA in Economics, University of Michigan	Private-sector banking executive (Continental Illinois); experience in commercial banking and financial operations.
Michael H. Moskow	1994–2007	BA in Economics, Lafayette College; PhD in Business & Applied Economics, University of Pennsylvania (Wharton)	Senior roles in government (Department of Labor, Department of Commerce); corporate executive; broad experience in labor markets and industrial policy.
Charles L. Evans	2007–2023	BA in Economics, University of Virginia; PhD in Economics, Carnegie Mellon University	Director of Research at Chicago Fed; academic economist; research on monetary policy, inflation dynamics, and macroeconomic modeling.
Austan D. Goolsbee	2023–present	BA & MA in Economics, Yale University; PhD in Economics, Massachusetts Institute of Technology	Professor of Economics at University of Chicago Booth School; former Chair of the Council of Economic Advisers; expert in empirical microeconomics and economic policy.

Table 14: Presidents of the Federal Reserve Bank of Chicago (1986–2025): Education and Professional Background

The apparent surge in speeches after the late 1990s is largely attributable to improvements in the Chicago Fed's archival completeness during that period. Once the archive stabilizes, speech frequency becomes noticeably higher during Michael H. Moskow's tenure (1994–2007), particularly in the years surrounding 2000. This coincides with several major economic events—the mid-1990s slowdown, the dot-com boom and bust, and Y2K preparations—which may have increased the demand for public communication. As in other districts, there is a downward shift toward a clearer, more accessible communication style after 2020.

Although abstractness during this period remains close to the district's long-run average, disunity and informativeness (Figures 27 and 28) decline, indicating that speeches became more focused on a narrower set of issues. The higher speech frequency also suggests that Moskow may

have delivered multiple specialized speeches, each addressing a specific topic, rather than fewer, broader speeches covering many themes at once. Readability difficulty and cognitive strain also fall during this period, mirroring the pattern in disunity and suggesting that Moskow's speeches were comparatively easier to follow. This is consistent with his background in business and applied economics, as well as his experience as a corporate executive and in industrial policy.

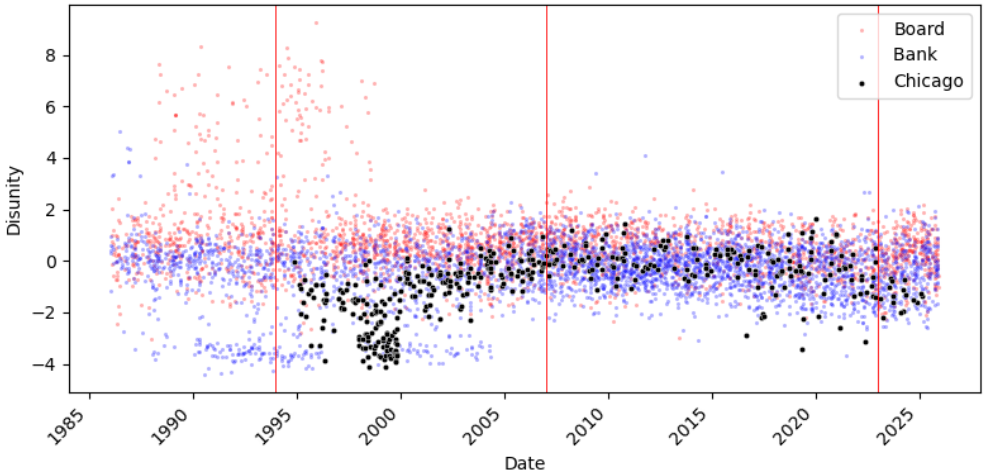

Figure 27: Federal Reserve Bank of Chicago Disunity

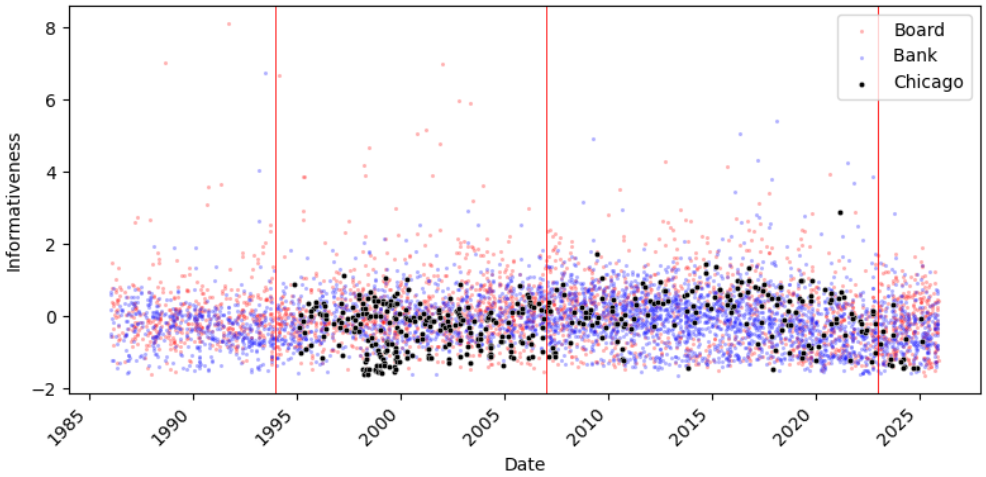

Figure 28: Federal Reserve Bank of Chicago Informativeness

7.3.7 Federal Reserve Bank of St. Louis

The president serving from 1986–1998 delivered somewhat fewer speeches than the presidents from 2008–2023 (and beyond, though the post-2023 sample remains small). Across most complexity

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Thomas C. Melzer	1985–1998	BS in Engineering, Princeton University; MBA (Finance), Stanford University	Private-sector finance executive (Morgan Stanley); expertise in financial markets, corporate finance, and banking operations.
William Poole	1998–2008	AB in Economics, Swarthmore College; PhD in Economics, University of Chicago	Senior economist at the Federal Reserve Board; professor at Brown University; research in monetary policy, macroeconomics, and financial markets.
James B. Bullard	2008–2023	BA in Economics & Quantitative Methods, St. Cloud State University; MA & PhD in Economics, Indiana University	Career Federal Reserve economist; vice president and deputy director of research; widely published macroeconomist; expertise in monetary theory and expectations.
Alberto G. Musalem	2024–present	BS & MS in Economics, London School of Economics; PhD in Economics, University of Pennsylvania	CEO and co-CIO of a quantitative investment firm; former IMF and private-sector finance expert; background in global markets, risk management, and financial technology.

Table 15: Presidents of the Federal Reserve Bank of St. Louis (1986–2025): Education and Professional Background

measures, St. Louis presidents fall close to both the district and cross-bank averages. Abstractness trends downward after 2020, consistent with the System-wide shift toward more practical and accessible communication during the pandemic period.

One notable feature is the pattern of informativeness during James Bullard's tenure (2008–2023), which is unusually low and remarkably stable, with very little variation across speeches. This is surprising given his professional background, as shown in Figure 29. By contrast, William Poole (1998–2008) exhibits consistently higher and rising complexity over his term, which aligns with his academic training and policy experience. The other presidents—Thomas C. Melzer, with an engineering/finance background, and Alberto Musalem, whose experience is primarily in private-sector leadership roles—display slightly lower complexity levels.

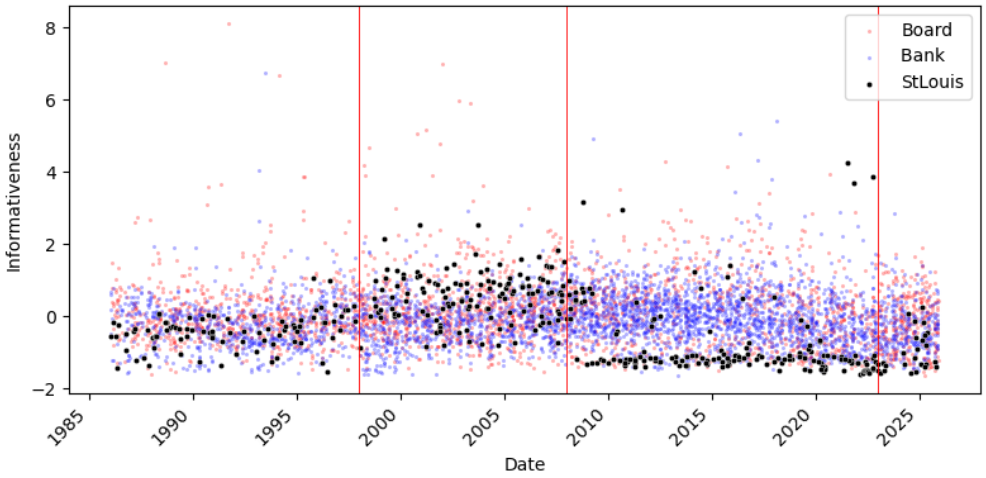

Figure 29: Federal Reserve Bank of St. Louis Informativeness

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Gary H. Stern	1985–2009	BA in Economics, Washington University in St. Louis; PhD in Economics, Rice University	Career Federal Reserve economist; Director of Research at Minneapolis Fed; co-author of *Too Big to Fail*; expertise in banking, financial stability, and macroprudential policy.
Narayana Kocherlakota	2009–2015	BA in Mathematics, Princeton University; PhD in Economics, University of Chicago	Academic economist (Stanford, Minnesota); research in monetary theory, macroeconomics, and expectations; former chair of University of Minnesota economics department.
Neel Kashkari	2016–present	BS & MS in Mechanical Engineering, University of Illinois Urbana–Champaign; MBA (Finance), University of Pennsylvania (Wharton)	Former Assistant Secretary of the Treasury; led TARP during the financial crisis; managing director at PIMCO; background in crisis management, public policy, and financial markets.

Table 16: Presidents of the Federal Reserve Bank of Minneapolis (1986–2025): Education and Professional Background

7.3.8 Federal Reserve Bank of Minneapolis

As with several other districts, the limited number of observations before the mid-2000s reflects archival gaps rather than true communication patterns. What is clear, however, is that Narayana Kocherlakota (2009–2016) delivered substantially more speeches than both his predecessor Gary

Stern (1985–2009) and his successor Neel Kashkari (2016–present). Because the number of speeches differs so sharply across these presidencies, direct comparisons of complexity metrics should be interpreted cautiously. Figures 30 and 31 show abstractness and cognitive strain for the Minneapolis Fed.

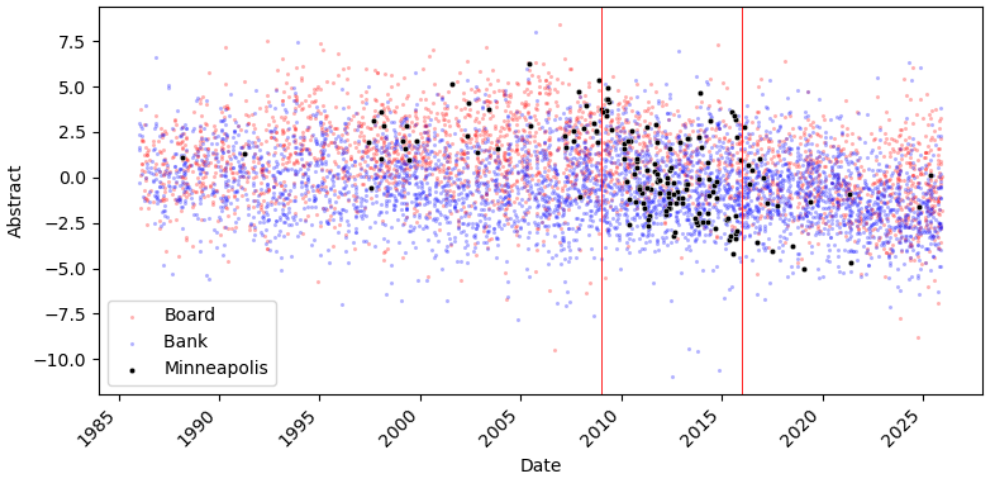

Figure 30: Federal Reserve Bank of Minneapolis Abstractness

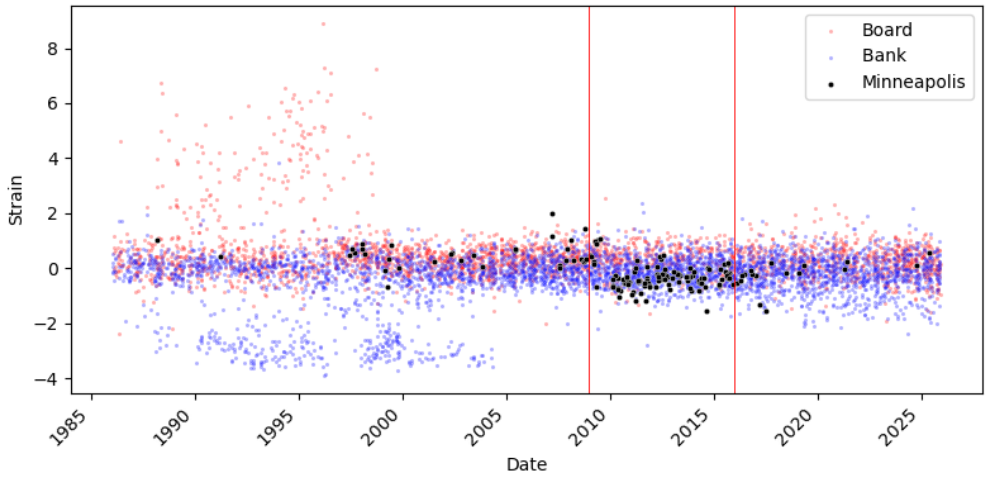

Figure 31: Federal Reserve Bank of Minneapolis Cognitive Strain

Overall, most complexity measures for Minneapolis presidents cluster around the district and cross-bank averages. The main exceptions are abstractness and cognitive strain (both versions), which are noticeably higher for Stern than for the later presidents. This aligns with Stern's background: his pre-2009 speeches often emphasized long-run structural issues and financial stability, consistent with his well-known focus on "too big to fail," culminating in his 2004 co-authored book

Too Big to Fail: The Hazards of Bank Bailouts.

Kocherlakota's tenure, by contrast, begins in the immediate aftermath of the Global Financial Crisis—a period marked by intense public communication, explanations of unconventional monetary policy, and efforts to rebuild credibility. He was also unusually transparent about his policy disagreements and evolving views, a feature noted in both his speeches and subsequent commentary. These factors likely contributed to his higher speech frequency and more policy-focused communication style.

7.3.9 Federal Reserve Bank of Kansas City

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
J. Roger Guffey	1976–1991	BA in Political Science, University of Missouri; JD, University of Missouri School of Law	Attorney; senior legal counsel at the Kansas City Fed; extensive experience in bank regulation and Federal Reserve operations.
Thomas M. Hoenig	1991–2011	BS in Economics, University of Missouri–Kansas City; PhD in Economics, University of Missouri	Career Federal Reserve economist; specialist in bank supervision and financial stability; known for strong views on monetary policy normalization.
Esther L. George	2011–2023	BS in Business Administration, Missouri Western State University; MBA, University of Missouri–Kansas City	Career Federal Reserve official; expert in bank supervision, community banking, and financial regulation; long tenure in Fed operations.
Jeffrey R. Schmid	2023–present	BBA in Finance & Economics, University of Nebraska–Lincoln; Banking Leadership Program, Southwestern Graduate School of Banking (SMU)	More than 40 years in banking and supervision; former FDIC field examiner; CEO of Mutual of Omaha Bank; president of the Southwestern Graduate School of Banking.

Table 17: Presidents of the Federal Reserve Bank of Kansas City (1986–2025): Education and Professional Background

Speeches are recorded only from 1995 onward, reflecting the point at which the Kansas City Fed's digital archive becomes consistently available. For both Thomas Hoenig (1991–2011) and Esther George (2011–2023), speech frequency increases toward the end of their respective tenures. This pattern may coincide with major macroeconomic events—the Global Financial Crisis

for Hoenig and the COVID-19 pandemic for George. During the pandemic period, abstractness (Figure 32) appears to decline after 2020, consistent with the broader System-wide shift toward more practical, operational communication during crisis conditions.

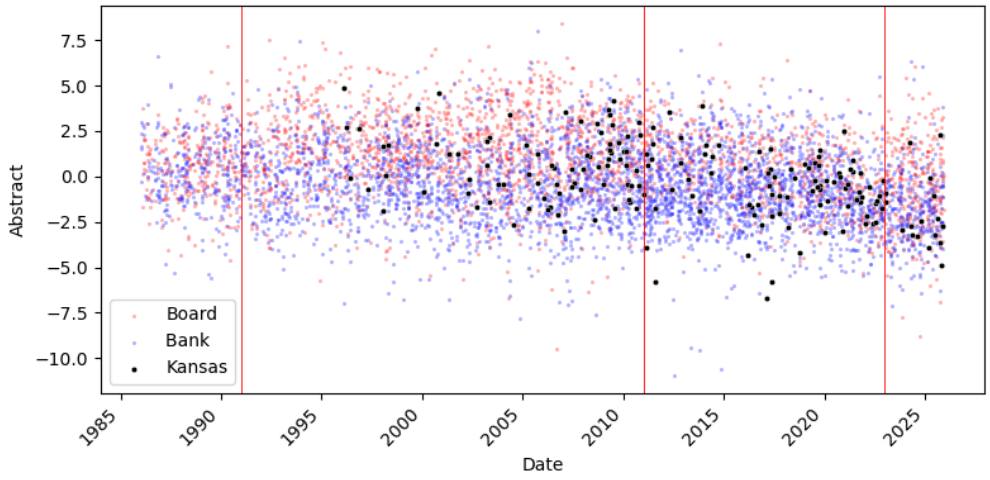

Figure 32: Federal Reserve Bank of Kansas Abstractness

For the remaining complexity measures—disunity, informativeness, readability difficulty, and cognitive strain—the values cluster closely around both the Kansas City average and the cross-bank average. There are no strong deviations or structural breaks. This raises the possibility that Kansas City presidents may exhibit a relatively consistent and stable communication style, with less variation across individuals than in some other districts. Although Hoenig and George come from different educational backgrounds—economics in Hoenig's case and business in George's—their long experience in banking supervision, financial stability, and operational policy may contribute to a more aligned communication style than their formal training alone would suggest.

7.3.10 Federal Reserve Bank of Dallas

Due to archival limitations, speeches are available only from around 1998 onward. The number of speeches delivered by Dallas Fed presidents is fairly consistent across most tenures, with the exception of Robert Kaplan (2015–2021), who gave relatively fewer speeches even early in his term, followed by a sharp decline near its end. This decline coincides with the period in 2021 when Kaplan became involved in a high-profile controversy regarding personal trading activity. While no public evidence confirms that he intentionally reduced public appearances in response, the timing suggests

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Robert H. Boykin	1981–1991	BA in Economics, Southern Methodist University; MBA, University of Texas at Austin	Career Federal Reserve official; extensive experience in bank supervision, financial operations, and regional economic analysis.
Robert D. McTeer	1991–2005	BA, MA, PhD in Economics, University of Georgia	Career Federal Reserve economist; president of Richmond Fed branch; known for accessible communication style and strong views on monetary policy and education.
Richard W. Fisher	2005–2015	BA in Economics, Harvard University; MBA, Stanford University; MA in Latin American Studies, Stanford University	Deputy U.S. Trade Representative; investment banker; expertise in international finance, trade policy, and global markets.
Robert S. Kaplan	2015–2021	BS in Business Administration, University of Kansas; MBA, Harvard Business School	Professor and Senior Associate Dean at Harvard Business School; former Goldman Sachs vice chairman; deep background in corporate finance and leadership.
Lorie K. Logan	2022–present	BA in Political Science, Davidson College; MA in Public Administration, Columbia University (SIPA)	Long career at the New York Fed; manager of the System Open Market Account (SOMA); expert in monetary policy implementation, financial markets, and liquidity operations.

Table 18: Presidents of the Federal Reserve Bank of Dallas (1986–2025): Education and Professional Background

that the episode may have contributed to a lower volume of speeches. Disunity and readability patterns are shown in Figures 33 and 34.

The speeches Kaplan did deliver tend to be more complex (with the exception of informativeness and the mean version of cognitive strain), which is consistent with his background: 23 years at Goldman Sachs and several years as a Harvard Business School professor and senior associate dean.

In terms of disunity, readability, and cognitive strain (both versions), Robert McTeer exhibits the lowest levels of complexity, followed by Richard Fisher and Lorie Logan. This aligns well with McTeer's background and reputation for maintaining a heavy speaking schedule, producing content accessible to general audiences, engaging frequently with the media, and incorporating poetry and

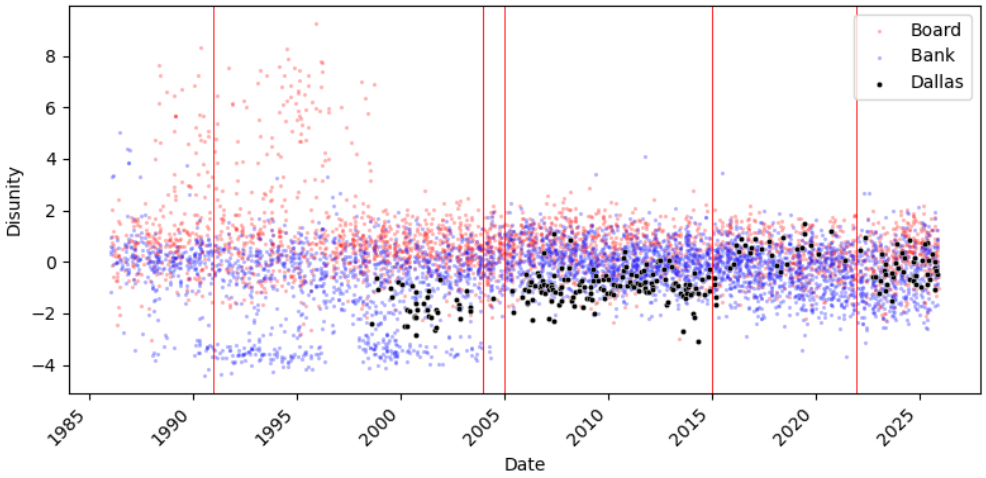

Figure 33: Federal Reserve Bank of Dallas Disunity

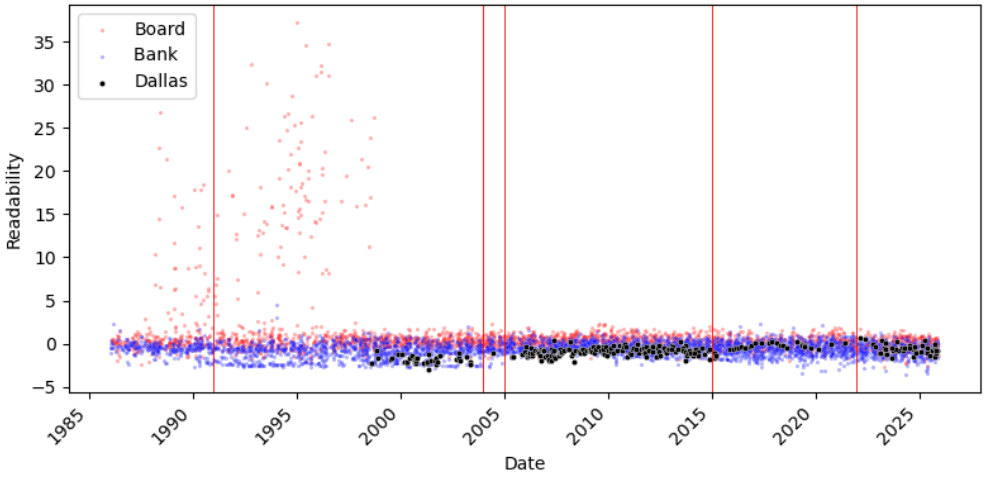

Figure 34: Federal Reserve Bank of Dallas Readability

vignettes into his communication. Fisher and Logan are relatively similar in complexity levels, with Fisher's speeches appearing somewhat clearer. This may reflect Logan's more technically demanding roles in market operations and implementation, which tend to produce less narrative-driven communication. Fisher, by contrast, combines an economics and business education with experience as Deputy U.S. Trade Representative, contributing to a communication style that is often colorful, direct, and rich in metaphors and anecdotes.

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Robert T. Parry	1986–2004	BA in Economics, University of California, Berkeley; PhD in Economics, University of Pennsylvania	Academic economist; professor at Wharton; senior economist at the Board of Governors; expertise in monetary policy and macroeconomic research.
Janet L. Yellen	2004–2010	BA in Economics, Brown University; PhD in Economics, Yale University	Professor at UC Berkeley; Chair of the Council of Economic Advisers; leading labor economist; later Chair of the Federal Reserve and U.S. Treasury Secretary.
John C. Williams	2011–2018	BA in Economics, University of California, Berkeley; MSc in Economics, London School of Economics; PhD in Economics, Stanford University	Research economist at the Fed; expert in monetary policy rules and macroeconomic modeling; later President of the New York Fed.
Mary C. Daly	2018–present	BA in Economics, University of Missouri–Kansas City; MA in Economics, University of Illinois Urbana–Champaign; PhD in Economics, Syracuse University	Career Federal Reserve economist; Executive VP and Director of Research at the San Francisco Fed; research in labor markets, inequality, and macroeconomics.

Table 19: Presidents of the Federal Reserve Bank of San Francisco (1986–2025): Education and Professional Background

7.3.11 Federal Reserve Bank of San Francisco

San Francisco has one of the richest and most complete speech archives in the Federal Reserve System, with records extending back to 1986. Speech frequency is consistently substantial throughout the sample, with the exception of a puzzling dip between roughly 1996 and 1998. Aside from this anomaly, the archive is continuous and highly consistent.

The complexity patterns for San Francisco are among the most distinctive in the dataset. Informativeness and cognitive strain for this district are shown in Figures 35 and 36.

Abstractness displays wide dispersion but tends to hover around the district and cross-bank averages. Disunity, however, is strikingly low during Robert Parry's presidency (1996–2004)—even lower than in most other districts. After 2005, disunity rises sharply and remains elevated, with a possible but subtle decline following Janet Yellen's tenure (2004–2010). Informativeness, readability difficulty, and cognitive strain follow a similar pattern to disunity, though their downward trend

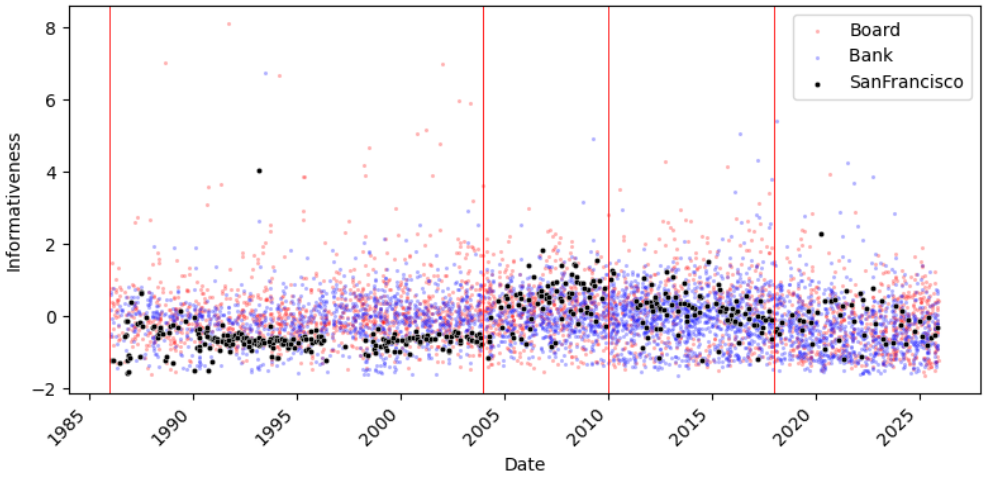

Figure 35: Federal Reserve Bank of San Francisco Informativeness

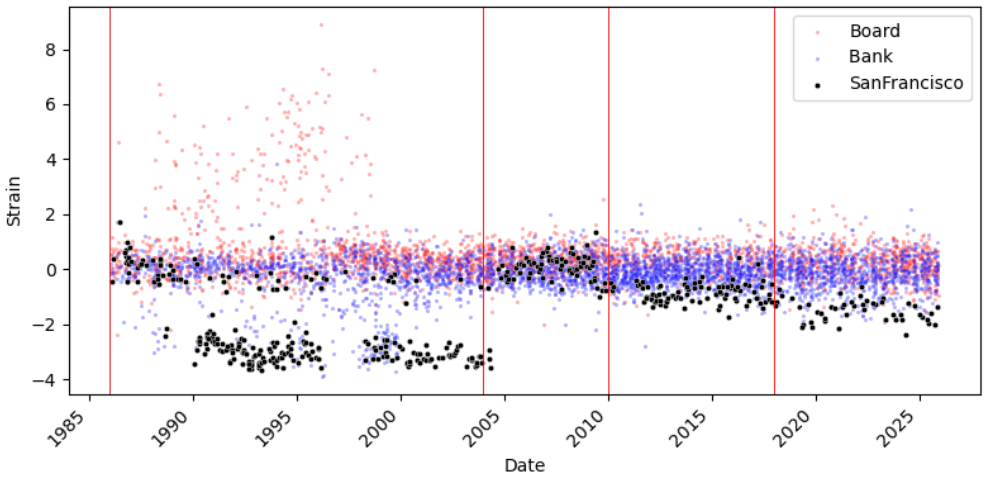

Figure 36: Federal Reserve Bank of San Francisco Cognitive Strain

in recent years is more pronounced. This pattern does not appear in the upper-bound version of cognitive strain, which remains relatively constant throughout.

Because the presidents in this district share broadly similar educational backgrounds, the differences in complexity appear to stem more from experience and communication choices than from formal training. Under Parry (1996–2004), vocabulary is simpler, more cohesive, and less structurally difficult, and his communication style appears relatively uniform. Under Yellen and her successors (2004 onward), speeches become more complex immediately during Yellen's tenure and then gradually decline again in the late 2010s and early 2020s. The decline is especially visible during the pandemic period, when several speeches appear unusually low in complexity—likely

reflecting highly practical, operationally focused communication.

7.3.12 Federal Reserve Bank of Atlanta

Name	Tenure	Education (Field + Institution)	Relevant Professional Background
Robert P. Forrestal	1983–1996	BA in English, St. John's University; JD (Law), Georgetown University Law Center	Attorney at the Federal Reserve Board; senior roles in Fed legal division; strong legal-institutional background influencing formal, structured communication.
Jack Guynn	1996–2006	BS in Industrial Management, Georgia Tech; MBA (Business Administration), University of Georgia	Career Federal Reserve official; expertise in regional economics, bank supervision, and payments systems; pragmatic, operations-oriented policymaker.
Dennis P. Lockhart	2007–2017	BA in Political Science, Stanford University; MA in International Economics & American Foreign Policy, Johns Hopkins SAIS	Managing director at Citigroup; business roles; deep experience in global finance and economic development.
Raphael W. Bostic	2017–present	BA in Economics & Psychology, Harvard University; PhD in Economics, Stanford University	Academic economist; professor of public policy at USC; research in housing, inequality, and governance; research-oriented communication style.

Table 20: Presidents of the Federal Reserve Bank of Atlanta (1986–2025): Education and Professional Background

Two presidents stand out for producing a notably large number of speeches: Robert Forrestal (1983–1996) and Dennis Lockhart (2007–2017). Despite their higher speech frequency, their overall communication complexity is broadly similar to that of other Atlanta presidents. There are slight differences in abstractness and cognitive strain (Figure 37); Lockhart's speeches appear marginally more concrete, though the difference is small. This may reflect his background, which is more oriented toward business and finance, whereas Forrestal came from a legal background.

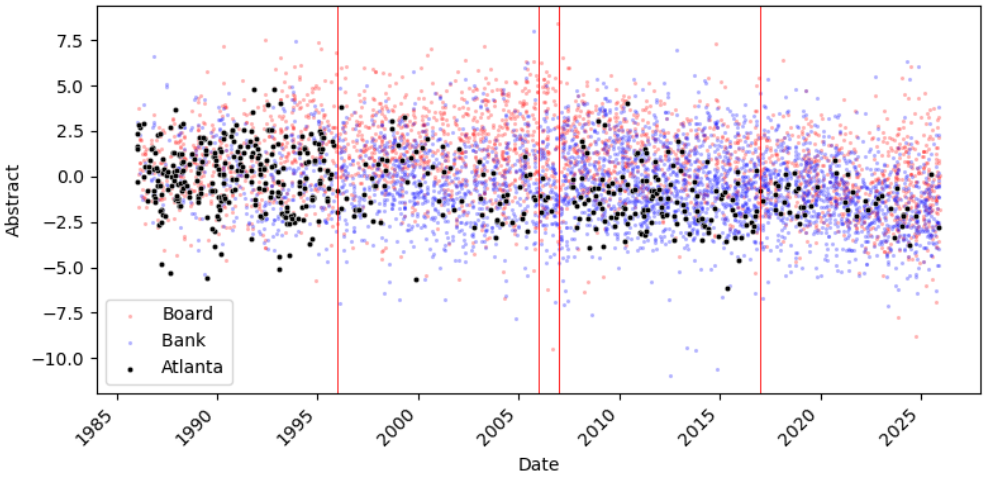

Figure 37: Federal Reserve Bank of Atlanta Abstractness

7.4 Construction of Grouped Topic Categories

To facilitate interpretation and reduce dimensionality, the twenty topics generated by the LDA model were subsequently aggregated into six broader thematic categories. This grouping is based on the semantic content of each topic's highest-probability keywords (reported in Appendix Table X). Although the LDA procedure identifies granular clusters of co-occurring terms, many of these topics naturally align with well-established domains of Federal Reserve communication. Collapsing the topics into a smaller number of conceptually coherent themes improves interpretability and avoids over-fitting when topic proportions are used as control variables in the empirical analysis.

The first category, Monetary Policy and Inflation, encompasses topics whose keywords relate to the conduct of monetary policy, inflation dynamics, and the FOMC's policy framework. Topics in this group include those emphasizing inflation, policy rules, interest rates, expectations, and the communication of the FOMC's long-run strategy (e.g., Topics 5, 7, 8, 9, and 17). These topics share a common focus on the Federal Reserve's core macroeconomic stabilization mandate and therefore form a unified thematic block.

A second category, Financial Stability and Banking Regulation, aggregates topics centered on bank supervision, capital regulation, systemic risk, crisis management, and credit market functioning. Topics in this group include those emphasizing supervisory processes, financial stability risks, liquidity stress, regulatory authority, and housing or mortgage credit (e.g., Topics 0, 1, 3, 16, and 19). These topics collectively reflect the Federal Reserve's role in safeguarding the financial

Topic	Top Keywords
0	reserve, rate, market, liquidity, _federal_reserve_, bank, security, facility, purchase, _balance_sheet_
1	bank, risk, institution, capital, supervisor, approach, can, supervisory, process, also
2	bank, _community_bank_, cra, community, datum, can, risk, firm, may, also
3	firm, crisis, risk, bank, institution, liquidity, _financial_system_, asset, capital, financial
4	market, risk, investor, asset, financial, can, price, may, value, _financial_market_
5	fed, policy, information, public, fomc, _central_bank_, _federal_reserve_, meeting, view, _monetary_policy_
6	economy, growth, year, recession, recovery, continue, spend, business, see, long
7	inflation, pandemic, see, datum, _labor_market_, economy, worker, high, supply, continue
8	fomc, long, policy, committee, low, see, expect, _monetary_policy_, rate, run
9	inflation, policy, _monetary_policy_, _central_bank_, long, can, fed, target, monetary, economy
10	productivity, increase, technology, economy, worker, business, new, industry, firm, high
11	increase, budget, tax, fiscal, debt, save, deficit, spend, income, benefit
12	country, u.s., _united_state_, trade, international, currency, foreign, market, world, dollar
13	think, go, know, one, say, get, just, person, can, good
14	payment, check, system, service, _payment_system_, bank, _federal_reserve_, use, new, currency
15	model, pp., vol., research, use, see, example, policy, may, estimate
16	bank, board, state, activity, regulation, act, require, provide, institution, banking
17	inflation, price, increase, rise, growth, long, rate, measure, forecast, expect
18	work, community, program, education, city, student, research, person, job, opportunity
19	loan, mortgage, credit, borrower, lender, lending, housing, bank, problem, rate

Table 21: LDA Topic–Keyword List (k = 20)

system and overseeing regulated institutions.

A third category, Real Economic Activity and Productivity, includes topics that describe macroeconomic conditions, business-cycle dynamics, productivity trends, and sectoral developments. Topics in this group emphasize growth, recession, employment, technological change, and the evolution of real economic indicators (e.g., Topics 6, 10, and 11). Because these topics jointly capture the Fed's assessment of underlying economic fundamentals, they are treated as a single thematic category.

A fourth category, Community Development and Inequality, groups topics associated with community outreach, local economic programs, education, and distributional concerns. These topics (e.g., Topics 2 and 18) are characterized by accessible, non-technical language and reflect

the Federal Reserve's public-facing communication on inclusive growth, community development initiatives, and the economic conditions of underserved populations.

A fifth category, International Economics and Global Spillovers, includes topics emphasizing international trade, exchange rates, global financial conditions, and cross-border policy interactions. Topic 12, for example, contains keywords related to foreign economies, currency markets, and global linkages. These topics share a global orientation and capture the Fed's discussion of external economic forces that influence domestic conditions.

Finally, a sixth category, Federal Reserve Operations and Payment Systems, aggregates topics related to the operational and institutional functions of the Federal Reserve System. Topics in this group include those emphasizing payment systems, check clearing, operational services, and the Fed's internal research and modeling infrastructure (e.g., Topics 14 and 15). These topics focus on the mechanics of financial infrastructure rather than on macroeconomic policy.

Together, these six thematic groups provide a parsimonious and economically coherent representation of the underlying topic structure. They allow the empirical analysis to control for systematic differences in speech content while avoiding the instability and multicollinearity that would arise from including all twenty granular topics individually. Because each group reflects a distinct domain of Federal Reserve communication, the aggregated categories also offer a natural framework for interpreting how different types of speech content relate to variation in linguistic complexity.

7.5 Financial Data as Predictors of Speech Complexity: Robustness Check

Category	Instrument	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Return	SPY	0	0	0	0	0	- -
	ES	0	0	0	0	0	0
	SHY	0	- -	+	++	0	+
	IEF	0	0	0	++	+	++
	TLT	0	0	0	++	0	+++
	UUP	0	+	- -	0	0	0
	VXX	0	0	0	0	0	++
	VIX	0	0	0	+	++	++
	GLD	0	0	0	0	0	0
Absolute Return	SPY	-	0	- - -	- - -	- - -	- - -
	ES	- -	0	+	+++	0	0
	SHY	0	0	0	0	- - -	-
	IEF	0	0	+	0	-	0
	TLT	0	0	+++	0	0	0
	UUP	-	0	0	0	0	0
	VXX	- - -	0	0	0	0	0
	VIX	0	0	0	+	0	0
	GLD	0	0	0	++	0	0
High-low Range	SPY	0	0	- - -	- - -	- - -	- - -
	ES	- -	0	++	++	0	0
	SHY	0	0	0	+	- -	-
	IEF	0	-	0	0	-	0
	TLT	0	-	+++	0	0	+
	UUP	0	0	0	++	0	- -
	VXX	- -	0	0	0	0	0
	VIX	0	0	+++	++	+	+++
	GLD	-	0	0	+++	0	0
Volume	SPY	- - -	- - -	- - -	- - -	- - -	0
	ES	-	0	0	0	0	+++
	SHY	- - -	- - -	++	0	0	0
	IEF	- - -	- - -	+++	- - -	0	+++
	TLT	- - -	- - -	0	- - -	0	+++
	UUP	0	0	- -	0	- - -	-
	VXX	- -	0	0	-	0	0
	GLD	- -	0	0	0	0	0
Abnormal Return (5 days)	SPY	0	0	0	0	0	- - -
	ES	0	0	0	0	0	-
	SHY	0	-	0	0	0	0
	IEF	0	0	0	+	0	0
	TLT	0	0	0	+	0	+
	UUP	0	0	0	0	0	0
	VXX	0	0	0	0	0	+
	VIX	- -	0	0	++	+	0
	GLD	0	0	0	0	0	0
Abnormal Return (22 days)	SPY	0	0	0	0	0	- -
	ES	0	0	+	0	0	-
	SHY	0	-	+	0	0	+
	IEF	0	0	0	+	0	++
	TLT	0	0	0	+	0	+++
	UUP	0	+	- -	0	0	0
	VXX	0	0	0	0	0	0
	VIX	-	0	0	0	+	0
	GLD	0	0	0	0	0	0

Table 22: Financial Data as Predictors of Speech Complexity with Signs

Notes: This table reports one-sided cumulative-effect tests of $H_0 : \beta_1+\beta_2+\beta_3+\beta_4 = 0$ for each asset–complexity pair in regression 1. Positive cumulative effects significant at the 1%, 5%, and 10% levels are denoted by "+++", "++", and "+"; negative cumulative effects by "- - -", "- -", and "-"; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include speaker fixed effects, macroeconomic controls, and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 3346 to 6468.

Category	Ticker	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Return	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	**	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	0
	UUP	0	0	0	0	**	0
	VXX	0	0	0	*	0	***
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0
Absolute Return	SPY	0	0	***	**	***	***
	ES	0	0	0	*	0	***
	SHY	0	0	0	0	*	0
	IEF	0	0	0	0	0	0
	TLT	*	0	0	0	0	0
	UUP	0	0	0	0	0	0
	VXX	0	**	0	0	0	*
	VIX	0	0	0	0	0	0
	GLD	0	0	**	0	**	0
High-low Range	SPY	***	0	***	***	***	***
	ES	0	0	0	0	0	*
	SHY	0	0	0	0	0	0
	IEF	**	0	*	0	0	0
	TLT	**	0	**	0	0	0
	UUP	*	0	0	0	0	0
	VXX	0	0	0	0	0	0
	VIX	0	0	**	0	0	***
	GLD	*	0	0	0	0	0
Volume	SPY	0	***	***	***	***	*
	ES	0	0	0	0	0	**
	SHY	**	**	0	0	0	0
	IEF	**	**	*	0	**	0
	TLT	**	0	0	0	0	0
	UUP	**	0	0	0	***	0
	VXX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0
Abnormal Return (5 days)	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	*	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	0
	UUP	0	0	0	0	**	0
	VXX	0	0	0	0	0	**
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	*	0
Abnormal Return (22 days)	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	**	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	*
	UUP	0	0	0	0	**	0
	VXX	0	0	0	0	0	***
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	*	0

Table 23: Financial Data as Predictors of Speech Complexity with Green/Tealbook

Notes: This table reports joint-significance tests of $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ for each asset–complexity pair in regression 1. Each cell presents the corresponding Wald test statistic and p-value. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include speaker fixed effects, macroeconomic controls (Greenbook/Tealbook), and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 3346 to 6468.

7.6 Speech Complexity as Predictors of Financial Data: Robustness Check

Institution	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Institution: New York	-0.26*	0.04	-0.73***	-0.20***	-0.18***	-0.33***
	(0.14)	(0.08)	(0.08)	(0.06)	(0.04)	(0.07)
Institution: Atlanta	-0.45***	-0.05	0.32***	0.21***	0.08	-0.61***
	(0.13)	(0.10)	(0.09)	(0.07)	(0.05)	(0.09)
Institution: San Francisco	-0.98***	-0.38***	-1.14***	-1.51***	-1.37***	-0.36***
	(0.12)	(0.06)	(0.08)	(0.08)	(0.06)	(0.06)
Institution: Chicago	-0.52***	-0.38***	-0.44***	-1.11***	-0.67***	-0.65***
	(0.15)	(0.08)	(0.10)	(0.09)	(0.07)	(0.08)
Institution: Cleveland	-0.47***	-0.13**	-0.12	-0.95***	-0.40***	-0.10
	(0.14)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)
Institution: Philadelphia	-0.23	-0.51***	-0.16*	-0.22***	-0.38***	-1.15***
	(0.19)	(0.16)	(0.09)	(0.08)	(0.06)	(0.13)
Institution: Richmond	-0.43***	-0.05	0.23***	0.59***	0.41***	-0.15
	(0.16)	(0.07)	(0.07)	(0.07)	(0.05)	(0.11)
Institution: StLouis	-1.07***	-0.44***	-0.72***	-0.50***	-0.40***	-0.71***
	(0.15)	(0.08)	(0.08)	(0.07)	(0.05)	(0.07)
Institution: Dallas	-0.69**	0.13	-0.77***	-0.52***	-0.61***	-0.34**
	(0.27)	(0.14)	(0.19)	(0.17)	(0.10)	(0.14)
Institution: Boston	-0.72***	0.06	-0.05	-0.38***	0.30***	0.31***
	(0.15)	(0.06)	(0.08)	(0.06)	(0.05)	(0.08)
Institution: Kansas	-0.39**	-0.41***	0.28**	-0.04	-0.13*	-0.43***
	(0.18)	(0.07)	(0.12)	(0.09)	(0.07)	(0.08)
Institution: Minneapolis	1.07***	-0.12	0.82***	0.48***	0.39***	-0.43***
	(0.21)	(0.15)	(0.16)	(0.12)	(0.10)	(0.13)
N	5726	5726	5726	5726	5726	5726
R^2_{adj}	0.31	0.117	0.245	0.411	0.4	0.258

Table 24: Financial Data as Predictors of Speech Complexity: Institution instead of Role

Notes: This table reports coefficient estimates from regression 1 using SPY trading volume as the financial variable. Each column corresponds to a different measure of textual complexity, and each row lists the associated coefficient estimate and Newey–West HAC standard error (Bartlett kernel, 5 lags) for the variable indicated on the left. All specifications include speakers' backgrounds with institution dummies, macroeconomic controls, and topic-proportion controls. The number of observations and R^2 for each specification are reported at the bottom of the table.

7.7 Robustness: Sub-sample Analysis

The main specification estimated on the sub-sample of Reserve Bank presidents does not differ meaningfully from the baseline results and is therefore omitted for brevity. This is unsurprising, given the relatively small number of non-presidential speeches in the dataset. Instead, this section focuses on the robustness check that excludes Chairman Greenspan's speeches, whose unusually high complexity—particularly prior to 1998—may distort the baseline estimates. Although the core conclusions remain intact, several notable deviations emerge.

First, in the first group of regressions (see Table 30), the predictive power of SPY volatility declines substantially across most complexity measures, whereas the results for other assets and volatility measures remain largely unchanged. Second, despite Greenspan's outlier status, the institutional-role results remain strong: Reserve Bank presidents continue to exhibit significantly lower speech complexity than Board Governors (see Table 31). Finally, while some coefficients linking speaker-background characteristics to complexity shift modestly, the overall patterns are stable.

Equity	Group	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
SPY	speaker	6.02%	5.70%	7.82%	17.88%	19.17%	11.21%
	finance	0.61%	0.16%	0.27%	0.21%	0.17%	0.29%
	macro	0.68%	0.28%	0.69%	0.43%	0.44%	0.36%
	topics	2.89%	1.77%	0.41%	9.83%	0.83%	3.16%
ES	speaker	7.83%	9.88%	20.52%	17.25%	27.54%	16.66%
	finance	0.44%	0.77%	0.1%	0.32%	0.39%	0.79%
	macro	0.95%	0.20%	0.38%	0.59%	0.80%	0.28%
	topics	3.42%	2.16%	36.11%	22.86%	2.18%	6.00%
SHY	speaker	7.95%	10.53%	19.33%	14.91%	24.35%	16.16%
	finance	0.61%	0.86%	0.32%	0.12%	0.48%	0.97%
	macro	0.65%	0.21%	0.29%	0.48%	0.81%	0.44%
	topics	3.82%	2.26%	37.55%	24.65%	2.30%	6.15%
IEF	speaker	8.11%	10.18%	19.44%	14.83%	24.05%	16.13%
	finance	1.1%	0.84%	0.32%	0.51%	0.39%	0.8%
	macro	0.61%	0.21%	0.34%	0.69%	0.83%	0.42%
	topics	3.78%	2.24%	37.59%	24.55%	2.32%	6.13%
TLT	speaker	7.87%	10.39%	19.32%	14.88%	23.98%	16.38%
	finance	1.0%	0.64%	0.22%	0.28%	0.26%	0.87%
	macro	0.74%	0.13%	0.27%	0.57%	0.82%	0.32%
	topics	3.80%	2.26%	37.88%	24.64%	2.40%	6.13%
UUP	speaker	5.58%	8.49%	16.84%	10.95%	21.41%	12.88%
	finance	0.87%	0.6%	0.26%	0.66%	0.56%	1.2%
	macro	1.02%	0.44%	0.73%	0.61%	0.77%	0.13%
	topics	5.05%	2.57%	44.21%	29.19%	3.91%	8.23%
VXX	speaker	19.40%	16.81%	17.41%	13.40%	26.96%	11.31%
	finance	4.5%	9.0%	0.87%	1.8%	2.6%	7.3%
	macro	8.79%	3.60%	0.12%	3.02%	2.16%	1.94%
	topics	3.28%	3.80%	52.72%	30.10%	9.82%	11.42%
VIX	speaker	6.17%	6.24%	9.79%	20.09%	22.05%	12.53%
	finance	0.28%	1.1%	0.46%	0.2%	0.31%	0.93%
	macro	0.52%	0.33%	1.14%	0.44%	0.64%	0.42%
	topics	3.22%	2.12%	0.83%	11.52%	1.19%	3.92%
GLD	speaker	7.52%	9.90%	18.32%	11.75%	20.46%	15.20%
	finance	0.37%	0.35%	0.27%	0.56%	0.48%	0.13%
	macro	1.51%	0.28%	0.50%	0.40%	0.56%	0.31%
	topics	4.58%	2.73%	39.67%	26.97%	3.59%	7.32%

Table 25: Variance decomposition of linguistic complexity using Greenbook/Tealbook data by speaker backgrounds, financial variables, macro variables and topic variables

This table reports the ANOVA decomposition for regression 1. Each row corresponds to a group of explanatory variables, and the columns report the associated sum of squares and share of explained variation for each complexity measure. The decomposition is based on the OLS specification with speaker-background variables, macroeconomic controls, financial variables, and topic-proportion controls. For the financial variable group, the decomposition aggregates all non-overlapping measures of each asset listed on the right-hand side.

The findings for the second group of regressions—where speech complexity predicts financial-market behaviour—are essentially unaffected by the exclusion of Chairman Greenspan.

Asset	Return	Absolute Return	High-low Range	Volume	Abnormal Return (5 days)	Abnormal Return (22 days)
SPY	0	- -	- -	++	0	0
ES	0	-	- -	-	0	0
SHY	0	- -	0	- - -	0	0
IEF	0	- - -	- - -	- - -	0	0
TLT	0	- - -	- - -	- - -	0	0
UUP	0	0	0	- -	0	0
VXX	0	-	- -	+	0	0
VIX	0	0	0		0	0
GLD	0	-	0	- - -	0	0

Table 26: Speech Complexity as Predictors of Financial Data with Sign

Notes: This table reports one-sided cumulative-effect tests of $H_0 : \beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6 = 0$ for each asset measure in regression 2. Positive cumulative effects significant at the 1%, 5%, and 10% levels are denoted by "+++", "++", and "+"; negative cumulative effects by "- - -", "- -", and "-"; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include macroeconomic controls (vintage ALFRED) and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 3346 to 6468.

Asset	Return	Absolute Return	High-low Range	Volume	Abnormal Return (5 days)	Abnormal Return (22 days)
SPY	0	***	***	***	0	0
ES	0	0	*	***	0	0
SHY	0	**	*	***	0	0
IEF	0	***	***	***	0	0
TLT	0	***	***	***	0	0
UUP	0	0	0	***	0	0
VXX	**	0	0	0	0	***
VIX	0	0	***		0	0
GLD	0	*	0	***	0	0

Table 27: Speech Complexity as Predictors of Financial Data Using Greenbook/Tealbook data

Notes: This table reports joint-significance tests of all complexity measures $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ for each asset measure in regression 2. Each cell presents the corresponding Wald test statistic and p-value. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include macroeconomic controls (Greenbook/Tealbook) and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 3346 to 6468.

	Return				Absolute Return				High-low Range				Volume			
	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m
SPY	0	0	0	0	0	0	0	+	0	0	0	- -	0	- - -	- -	- - -
SHY	0	0	–	-	0	0	0	0	0	0	+	0	0	- -	- -	- - -
IEF	0	0	- - -	0	+++	++	++	++	++	0	0	0	0	- - -	- -	- -
TLT	0	0	- -	0	+++	++	0	0	+++	++	++	0	0	- -	0	- - -
VIX	0	0	0	0	0	0	0	0	0	0	0	0				

Table 28: Speech Complexity as Predictors of Financial Data with Intraday Data and Signs: Post-speech Reaction

Notes: This table reports one-sided cumulative-effect tests of $H_0 : \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 = 0$ for each asset measure in regression 2. Each column corresponds to a specific intraday window for a given asset measure. Positive cumulative effects significant at the 1%, 5%, and 10% levels are denoted by "+++", "++", and "+"; negative cumulative effects by "- - -", "- -", and "-"; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include macroeconomic controls (vintage ALFRED) and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 787 to 847.

	Return				Absolute Return				High-low Range				Volume			
	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m
SPY	0	0	0	0	0	0	0	0	0	0	0	***	**	***	***	***
SHY	0	0	0	0	**	0	0	0	0	0	0	0	0	**	0	**
IEF	0	0	0	0	0	*	0	0	***	0	0	0	*	0	*	**
TLT	0	0	0	0	*	0	0	0	*	0	0	0	0	**	0	***
VIX	0	0	0	0	0	0	0	0	0	0	0	***				

Table 29: Speech Complexity as Predictors of Financial Data with Intraday Data: Around-speech Reaction

Notes: This table reports joint-significance tests of all complexity measures $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ for each asset measure in regression 3. Each column corresponds to a specific intraday window for a given asset measure, and each cell presents the corresponding Wald test statistic and p-value. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively; "0" indicates no significance. All regressions are estimated by OLS with Newey–West HAC standard errors (Bartlett kernel, 5 lags) and include macroeconomic controls (vintage ALFRED) and topic proportions. Sample sizes differ across asset measures due to varying data availability; N ranges from 787 to 847.

Category	Instrument	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Return	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	**	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	0
	UUP	0	0	0	0	*	0
	VXX	0	0	0	0	0	**
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0
Absolute Return	SPY	0	0	0	0	0	**
	ES	0	0	0	**	0	**
	SHY	0	0	0	0	**	0
	IEF	**	0	0	0	0	0
	TLT	***	0	**	0	0	0
	UUP	0	0	0	0	0	0
	VXX	**	**	0	0	0	*
	VIX	0	0	0	0	0	0
	GLD	0	0	***	*	*	0
High-low Range	SPY	***	0	0	0	0	0
	ES	**	0	0	0	0	0
	SHY	0	0	0	0	0	0
	IEF	**	0	**	0	0	0
	TLT	**	0	***	0	0	0
	UUP	*	0	0	0	0	0
	VXX	0	0	0	0	0	0
	VIX	0	*	0	0	0	**
	GLD	**	0	0	***	0	0
Volume	SPY	***	***	**	0	0	0
	ES	0	0	0	*	0	***
	SHY	***	***	0	0	0	0
	IEF	***	***	**	***	0	**
	TLT	***	***	0	***	0	***
	UUP	**	0	0	0	***	0
	VXX	0	0	**	0	0	0
	GLD	*	0	0	0	0	0
Abnormal Return (5 days)	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	*	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	0
	UUP	0	0	0	0	0	0
	VXX	**	0	0	0	0	0
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0
Abnormal Return (22 days)	SPY	0	0	0	0	0	0
	ES	0	0	0	0	0	0
	SHY	0	*	0	0	0	0
	IEF	0	0	0	0	0	0
	TLT	0	0	0	0	0	*
	UUP	0	0	*	0	*	0
	VXX	0	0	0	0	0	0
	VIX	0	0	0	0	0	0
	GLD	0	0	0	0	0	0

Table 30: Financial Data as Predictors of Speech Complexity: Excluding Chairman Greenspan

Variable	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
Gender	-0.34***	-0.22***	-0.20***	-0.15***	-0.22***	0.05
	(0.08)	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)
Race	0.26*	0.06	0.37***	0.34***	0.18***	0.49***
	(0.14)	(0.06)	(0.04)	(0.06)	(0.05)	(0.05)
Major: Law	0.19	0.10	-0.12**	-0.25***	-0.00	0.11*
	(0.16)	(0.07)	(0.05)	(0.08)	(0.06)	(0.06)
Major: Business	-0.54***	-0.04	-0.08*	0.07	0.03	0.17***
	(0.15)	(0.07)	(0.04)	(0.08)	(0.06)	(0.06)
Major: Finance	-1.99***	-0.54**	-0.45***	0.32**	0.10	-0.11
	(0.36)	(0.27)	(0.13)	(0.13)	(0.11)	(0.24)
Major: Other	-0.33***	0.01	-0.11***	-0.06	0.11***	-0.18**
	(0.10)	(0.10)	(0.03)	(0.05)	(0.04)	(0.08)
Not PhD	-0.00	-0.10*	0.06**	-0.04	-0.14***	-0.32***
	(0.09)	(0.06)	(0.03)	(0.06)	(0.05)	(0.05)
Pre-Fed: Academic	0.45***	0.17***	0.14***	0.19***	0.17***	-0.01
	(0.08)	(0.05)	(0.02)	(0.04)	(0.03)	(0.04)
Pre-Fed: Finance	0.21**	0.02	0.17***	0.38***	0.26***	-0.12*
	(0.10)	(0.09)	(0.03)	(0.05)	(0.04)	(0.07)
Pre-Fed: Business	-0.88***	0.51***	0.40***	0.22**	0.50***	0.35***
	(0.25)	(0.11)	(0.07)	(0.11)	(0.08)	(0.10)
Pre-Fed: Law	-0.43**	-0.15	0.05	0.40***	0.42***	-0.07
	(0.19)	(0.10)	(0.06)	(0.10)	(0.07)	(0.09)
Pre-Fed: Other	0.23	0.04	0.32***	0.48***	0.23***	0.22***
	(0.16)	(0.07)	(0.06)	(0.09)	(0.05)	(0.08)
Freshwater	0.18*	0.19***	0.14***	0.28***	0.17***	-0.13***
	(0.10)	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)
Other Economist	-0.27***	-0.26***	0.09***	0.25***	0.19***	-0.25***
	(0.10)	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)
Age	-0.02***	0.01***	-0.00	-0.00	0.00*	0.01**
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Terms Time	0.03***	-0.02***	0.00	-0.02***	-0.01**	0.01*
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
gINDPRO	4.38	-3.86*	1.79	3.81*	1.53	-2.66
	(4.10)	(2.07)	(1.30)	(2.12)	(1.62)	(1.99)
gINDPRO lag1	1.41	0.45	2.06***	0.82	1.04	1.77*
	(2.56)	(1.11)	(0.75)	(1.12)	(0.92)	(1.03)
Infl	5.03	0.11	-0.41	-6.19	-2.00	-3.40
	(7.03)	(2.67)	(2.05)	(3.82)	(2.65)	(2.86)
Infl lag1	-3.59	-2.11	-1.44	0.32	0.37	-0.63
	(6.43)	(2.46)	(1.87)	(3.52)	(2.42)	(2.51)
MICH	-0.14	-0.01	0.02	0.04	0.03	0.07**
	(0.08)	(0.04)	(0.02)	(0.04)	(0.03)	(0.04)
MICH lag1	-0.13	-0.03	-0.01	0.03	-0.02	-0.02
	(0.09)	(0.04)	(0.03)	(0.05)	(0.03)	(0.04)
Unemp	0.05	-0.09***	0.01	0.00	-0.00	-0.04*
	(0.07)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Unemp lag1	0.01	0.09***	-0.01	-0.02	0.00	0.04*
	(0.07)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Topic: Community Dev	-0.27	-0.78***	2.12***	1.74***	0.56***	0.77***
	(0.32)	(0.16)	(0.11)	(0.14)	(0.10)	(0.12)
Topic: MonPol–Inflation	0.47*	0.31**	1.93***	2.49***	0.93***	0.91***
	(0.28)	(0.14)	(0.09)	(0.12)	(0.09)	(0.09)
Topic: Real Econ–Prod	-1.13***	0.20	0.78***	1.20***	-0.19*	0.43***
	(0.29)	(0.14)	(0.09)	(0.14)	(0.10)	(0.10)
Topic: Fed Ops–Payments	1.00***	0.91***	2.03***	2.72***	0.81***	2.46***
	(0.37)	(0.21)	(0.11)	(0.15)	(0.12)	(0.25)
Topic: FinStab–Reg	3.05***	0.52***	2.80***	2.82***	1.01***	0.81***
	(0.30)	(0.16)	(0.09)	(0.13)	(0.09)	(0.10)
Topic: International Econ	2.77***	0.51***	2.70***	3.53***	1.48***	0.94***
	(0.35)	(0.16)	(0.11)	(0.17)	(0.12)	(0.14)
Is Bank Pres	-0.43***	-0.23***	-0.43***	-0.44***	-0.36***	-0.45***
	(0.07)	(0.05)	(0.02)	(0.04)	(0.03)	(0.04)
N	5253	5253	5253	5253	5253	5253
R^2_{adj}	0.296	0.111	0.568	0.323	0.282	0.24

Table 31: Financial Data as Predictors of Speech Complexity Excluding Chairman Greenspan: Speaker background, macroeconomic controls, and topics

Equity	Group	Abstract	Informativeness	Readability	Disunity	Strain	Strain (upper)
SPY	speaker	4.19%	5.97%	20.12%	18.64%	24.27%	13.95%
	finance	0.72%	0.14%	0.14%	0.22%	0.21%	0.35%
	macro	0.56%	0.55%	0.29%	0.26%	0.21%	0.53%
	topics	5.34%	1.84%	36.81%	22.48%	3.64%	7.11%
ES	speaker	6.83%	11.16%	20.23%	15.46%	28.99%	18.11%
	finance	0.51%	0.9%	0.13%	0.32%	0.29%	0.72%
	macro	0.96%	1.04%	0.49%	0.28%	0.71%	1.00%
	topics	4.72%	2.06%	39.56%	25.07%	2.67%	6.44%
SHY	speaker	6.39%	12.13%	18.88%	12.76%	25.86%	17.87%
	finance	1.1%	0.82%	0.4%	0.47%	0.6%	0.82%
	macro	1.09%	0.68%	0.62%	0.41%	0.88%	1.10%
	topics	4.80%	2.15%	40.86%	27.33%	3.11%	6.71%
IEF	speaker	6.91%	11.81%	19.27%	12.51%	25.70%	18.14%
	finance	1.5%	0.57%	0.44%	0.82%	0.39%	0.6%
	macro	0.81%	0.58%	0.64%	0.47%	0.85%	0.82%
	topics	4.60%	2.17%	40.61%	27.12%	3.07%	6.59%
TLT	speaker	6.60%	12.05%	18.87%	12.29%	25.53%	18.29%
	finance	1.3%	0.65%	0.22%	0.41%	0.14%	0.78%
	macro	1.09%	0.59%	0.56%	0.36%	0.69%	0.80%
	topics	4.75%	2.18%	41.06%	27.37%	3.11%	6.47%
UUP	speaker	4.81%	8.59%	14.64%	7.86%	19.67%	13.44%
	finance	0.86%	0.73%	0.29%	0.41%	0.48%	0.8%
	macro	0.82%	0.73%	0.49%	0.35%	0.93%	0.79%
	topics	5.77%	2.43%	49.35%	34.15%	6.30%	8.37%
VXX	speaker	19.16%	16.52%	9.31%	6.96%	13.29%	2.95%
	finance	3.4%	3.2%	0.82%	0.72%	1.9%	2.3%
	macro	3.08%	1.38%	1.81%	0.54%	3.22%	2.04%
	topics	4.03%	3.28%	59.92%	35.86%	10.33%	11.26%
VIX	speaker	5.26%	7.16%	22.03%	21.05%	27.65%	16.31%
	finance	1.1%	0.77%	0.19%	0.18%	0.22%	0.57%
	macro	0.85%	0.70%	0.42%	0.38%	0.52%	1.03%
	topics	5.48%	2.14%	36.19%	21.44%	2.87%	7.71%
GLD	speaker	6.31%	10.49%	17.00%	9.65%	21.11%	16.68%
	finance	0.65%	0.22%	0.44%	0.71%	0.54%	0.11%
	macro	1.07%	0.72%	0.67%	0.36%	0.85%	1.18%
	topics	5.44%	2.49%	43.67%	30.52%	5.08%	7.31%

Table 32: Variance decomposition of linguistic complexity by speaker backgrounds, financial variables, macro variables and topic variables, excluding Chairman Greenspand

Asset	Return	Absolute Return	High-low Range	Volume	Abnormal Return (5 days)	Abnormal Return (22 days)
SPY	0	***	***	***	0	0
ES	0	0	0	***	0	0
SHY	0	***	***	***	0	0
IEF	*	***	**	***	0	0
TLT	0	***	***	***	0	0
UUP	0	0	**	0	0	0
VXX	0	0	*	***	0	0
VIX	0	*	***		0	0
GLD	0	0	*	***	0	0

Table 33: Speech Complexity as Predictors of Financial Data: return, volatility, and abnormal return measures across assets, excluding Chairman Greenspan

	Return				Absolute Return				High-low Range				Volume			
	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m	15m	30m	45m	60m
SPY	0	0	0	0	0	0	***	0	0	0	*	**	0	***	**	***
SHY	0	0	*	0	*	0	0	0	**	0	0	0	0	**	0	**
IEF	0	0	0	0	***	0	0	*	**	0	**	0	0	***	**	***
TLT	0	0	0	0	*	0	0	0	*	0	**	0	0	**	0	***
VIX	0	0	0	0	0	0	0	*	0	**	0	0				

Table 34: Speech Complexity as Predictors of Financial Data with Intraday Data: Post-speech Reaction, excluding Chairman Greenspan