Research-Thesis: Central Bank Communication Over The Last Decade: Has Forward Guidance Become More Odyssean?

Auteur : Debra, Benjamin
Promoteur(s) : Clerc, Pierrick
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences de gestion, à finalité spécialisée en Banking and Asset Management
Année académique : 2025-2026
URI/URL : http://hdl.handle.net/2268.2/25213

CENTRAL BANK COMMUNICATION OVER THE LAST DECADE: HAS FORWARD GUIDANCE BECOME MORE ODYSSEAN ?

Jury :
Supervisor :
Dr. Pierrick Clerc
Reader :
Romain Crucil

Master thesis presented by
Benjamin Debra
To obtain the degree of Master in Management with a specialization in
Banking and Asset Management
Academic year 2025-2026

Ecole de Gestion de l'Université de Liège

Acknowledgements

First of all, I would like to express my sincere gratitude to my thesis supervisor, Professor Pierrick Clerc, for proposing such a stimulating and ambitious research topic. His valuable advice and recommendations were a great help to me throughout the writing of this thesis.

I would also like to thank my first thesis lecturer, Romain Crucil, for his constant availability and precious help during the preparation of this work. His assistance has been greatly appreciated.

I would like to thank my family, especially my mother, whose support made my academic journey possible. I would also like to thank all those who contributed to my personal and intellectual development through their knowledge, advice, and support during difficult times.

Finally, I would like to give special thanks to my close friends PA, Jérôme, Felix, Matthieu, Damchi, Martin, Grue, and Matthias for their optimism and the countless memorable moments we shared during these five years at university. I am equally grateful for all the wonderful people I have met , who have made this experience truly enriching.

Table of Contents

List of abbreviations

FED : Federal Reserve System
ZLB/ELB : Zero/Effective Lower Bound
NLTK : Natural Language Toolkit
LM : Loughran and McDonald
FG : Forward Guidance
OLS : Ordinary Least Square
VIF : Variance Inflation Factors
NLP : Natural Language Processing

1. Introduction

Over the last decade, the landscape of monetary policy has undergone a fundamental transformation where "words" have become highly influential alongside interest-rate decisions. This shift became particularly noticeable when central banks reached the "effective lower bound" (ELB) or "zero lower bound" (ZLB), a state where conventional rate cuts are no longer a viable policy option (Singh, 2023). In this constrained environment, communication is no longer merely supplementary; it became a central component of the monetary policy itself (Blinder et al., 2008). Central banks use forward guidance (FG) to signal future policy intentions, aiming to shape market expectations and stabilize financial conditions by influencing the entire yield curve (Moessner et al., 2015).

A central pillar of the modern discourse on forward guidance is the theoretical distinction between two types of signals: "Odyssean" guidance, which represents a binding commitment to a future policy stance, and "Delphic" guidance, which provides conditional forecasts and information about the macroeconomic outlook (Campbell et al., 2012). Academic debate often treats these as distinct dimensions that can coexist within the same announcement (Moessner et al., 2015). Despite this rich theoretical framework, a significant "disconnect" persists between theory and practice, as policymakers often prioritize flexibility and discretion over the rigid commitments suggested by theoretical models (Moessner et al., 2015).

This thesis addresses a research gap by developing a quantitative research design framework to measure these tones directly from central bank communication. While various channels exist for policy disclosure, this study focuses specifically on speeches. Speeches serve as a critical channel for clarifying policy intentions, offering a more richer and frequent medium than formal committee statements or press conferences (Swanson & Jayawickrema, 2024). Individual speeches are particularly informative because, in a context where central bank communication is dispersed among several policymakers, they offer a clear way to track the evolution of the institution's messages over time. Accordingly, this thesis seeks to answer the following central research question: « Has forward guidance become more Odyssean over the last decade » and will examine how this evolution manifests across two different central banking frameworks.

To explore the diversity of communication strategies, the thesis conducts a comparative analysis between the U.S. Federal Reserve and the Sveriges Riksbank. These two institutions represent divergent philosophies: the Riksbank is a pioneer of Delphic transparency, frequently publishing explicit interest-rate paths (Andersson & Hofmann, 2009), whereas the Federal Reserve has historically used a more episodic approach to guidance (Moessner et al., 2015). By examining these banks over the 2012–2024 period, the thesis captures their responses to varied macroeconomic regimes, including the stresses of the Covid -19 pandemic.

The study employs sentiment analysis to convert the qualitative content of these speeches into quantitative indices. Using a lexicon-based approach grounded in finance-specific dictionaries (Loughran & McDonald, 2011), the thesis measures the prevalence of Odyssean and Delphic language across the corpus. These tone indices are then used in an econometric framework to evaluate the relationship between communication style and government bond market volatility, focusing on short-to-medium-term maturities that are most sensitive to monetary policy expectations (Swanson, 2017).

The remaining sections of the thesis is structured as follows. Chapter 2 provides a detailed literature review of central bank communication, the Odyssean–Delphic dichotomy, and the historical evolution of the Fed and the Riksbank. Chapter 3 outlines the methodology, including the web-scraping process,

the text-cleaning pipeline, and the construction of the sentiment indices. Chapter 4 presents the results and regression analysis. Chapter 5 discusses the findings in the context of policy coordination and design, and Chapter 6 concludes with a summary and recommendations for future research.

2. Literature review

2.1 Context

During significant financial tensions or recessions, conventional monetary policy tools may become ineffective. This is particularly true in an "effective lower bound" (ELB)/ « Zero-lower bound » (ZLB) situation, where interest rates are at or near zero, rendering further rate cuts impossible (Singh, 2023). In such scenarios, central banks turn to unconventional monetary policy (Swanson, 2017 ; Singh, 2023). The global financial crisis, for instance, necessitated a rapid adoption of such tools (Assenmacher and al., 2021). These include quantitative easing (QE) and forward guidance (FG) (Singh,2023). Foward guidance is a communication-based tool that provides signals about future monetary policy actions to clarify uncertainty and influence market expectations (Singh,2023). In these ELB/ZLB contexts, communication is not merely supplementary; it becomes the "essence" of monetary policy itself (Bernanke et al., 2004; Eggertsson & Woodford, 2003, cited in Blinder et al., 2008).

This necessary shift toward unconventional tools highlights a broader, dramatic evolution in central bank communication. For a large portion of the 20th century, a "culture of silence and secrecy" predominated (Assenmacher et al., 2021 ; Masciandaro et al., 2024). This contrasts severely with the modern view. In recent decades, communication has transformed into a key instrument of openness and transparency (Geraats, 2002; Demertzis & Hallett, 2007; Blinder et al., 2008; Dincer & Eichengreen, 2014; cited in Masciandaro et al., 2024). It is now considered a fundamental tool of monetary policy (Bernanke, 2004; Blinder et al., 2008), and a fundamental mechanism that directly influences public perceptions regarding consumption (Assenmacher et al., 2021).

This increased transparency, however, is not without its challenges. Both Coenen et al. (2017) and Assenmacher et al. (2021) caution that by being more transparent, central banks risk generating "noise," if the information provided to the public about monetary policy is misunderstood and therefore increases market volatility. If this noise is not managed, it can dramatically impair the effectiveness of policy tools (Assenmacher et al., 2021). Consequently, the optimal communication strategy is not necessarily one of maximum disclosure. Blinder et al. (2008) advocate for maximizing the "signal-to-noise" ratio, which requires better-designed communication that is transparent about genuine intentions and reduces market uncertainty by minimizing the « noise » (cited in Assenmacher et al., 2021).

2.2 Forward Guidance

Forward guidance, broadly defined by Coenen et al. (2017) as "communication about the future conduct of monetary policy," is not a new concept, but its prominence and function were fundamentally transformed at the ZLB. The literature of this period materialized around a new consensus: policy efficacy in a liquidity trap is less dependent on current policy actions and almost entirely reliant on the central bank's ability to credibly manage public and market expectations about the future path of interest rates (Krugman ,1998 ; Woodford,2003 ; cited by Moessner et al.2015). By shaping expectations, a central bank can influence the entire yield curve (Moessner et al., 2015). This elevation of communication from a supplementary tool to a primary, key policy instrument forced a rigorous academic and practical debate on how such guidance should be designed, communicated, and interpreted.

2.2.1 The Core theoretical distinction: Delphic guidance vs. Odyssean guidance

The extensive literature that came out to analyze forward guidance is dominated by a central theoretical dichotomy, first defined by Campbell and al. (2012), between "Delphic" and "Odyssean" forms of communication (Moessner et al., 2015). This distinction is not merely semantic; it represents two fundamentally different mechanisms of policy transmission, which have different implications for economic stability and the central bank's credibility.

Delphic forward guidance, named for the Oracle of Delphi, is a purely informational act. In this mode, the central bank publishes its forecasts for the macroeconomy (inflation, output) and its likely monetary policy response given that outlook, but it makes no binding promise to follow that path (Moessner et al., 2015; Goy et al., 2018). As described by Coenen et al. (2017), the central bank "reserves the right to re-optimize its plan in every future period." Its effectiveness, therefore, depends entirely on the public perceiving the central bank as having superior information, better forecasting models, or, as Bassetto (2019) argues, private information about its own preferences and beliefs.

In theoretical opposition stands the Odyssean forward guidance. This form is not a forecast but a binding commitment (Coenen et al., 2017). The name invokes the myth of Odysseus, who, to resist the sirens' call, "ties himself to the mast" (Moessner et al., 2015). In this policy framework, the central bank explicitly promises a future course of action,such as keeping interest rates at zero for a prolonged period,and pledges to stick to that path even when future economic conditions might tempt it to deviate.The primary framework for this approach was laid by Eggertsson and Woodford (2003) in their analysis of the ZLB. They postulated that the only optimal path out of a liquidity trap is a history-dependent policy rule. This strategy requires a promise of a future inflation overshoot to "make up" for past weakness. The problem is that this promise is time-inconsistent, as the bank will be tempted to break it later. However, if the commitment is credible now, it powerfully stimulates the economy by lowering real rates via higher inflation expectations. (Eggertsson & Woodford, 2003)

This theoretical ideal was forcefully advocated by Michael Woodford (2012) during the crisis, who argued that for guidance to be successful, it must involve an "advance commitment to definite criteria for future policy decisions" (Woodford, 2012), branding mere forecasts as "insufficient." He proposed practical, history-dependent implementations such as a "gap-adjusted price level target" or, as a more easily understood alternative, nominal GDP target path (Woodford, 2012). This Odyssean approach is posited as the theoretically ideal monetary policy at the ZLB, capable of fully neutralizing a recession if the commitment is believed (Eggertsson & Woodford, 2003; McKay et al., 2016).

2.2.2 The « Forward Guidance Puzzle »

Even though the Odyssean framework was theoretically ideal, it was instantly hit by a major empirical contradiction, which is known as the "forward guidance puzzle" (Del Negro et al., 2012).This "Puzzle" identifies a significant disconnect: standard New Keynesian models predict that Odyssean promises about the future have unrealistically large and even counterintuitive effects on macroeconomic variables (Del Negro et al. 2012 ; cited in Hagedorn et al., 2019). Carlstrom et al. (2012), for instance, found that a standard model would predict "explosive inflation and output" from such a policy (cited in Moessner et al., 2015). Del Negro et al. (2012) confirm that these models "grossly overestimate" the impact. A wave of research has been conducted to resolve this contradiction by questioning the fundamental assumptions of the models, as this theoretical potency contrasts sharply with modest, empirical effects in the real world.

The literature has proposed several distinct resolutions. The first, and perhaps most influential resolution comes from introducing incomplete markets, a correction supported by McKay, Nakamura, and Steinsson in 2016. They argue that the standard model's assumption of "complete markets" (where households are perfectly insured against every possible future risk and have no borrowing constraints) is the source of the anomaly (McKay et al.2016). In their more realistic model with heterogeneous agents who face uninsurable income risk and borrowing constraints, households have a strong precautionary savings motive. This makes them far less responsive to promises about interest rates far in the future. This friction "substantially attenuates" the power of forward guidance, bringing its predicted impact more in line with modest, observed effects. Hagedorn et al. (2019) reach a similar conclusion in their incomplete markets model, finding that the intertemporal substitution channel is so dampened that FG effects become "negligible."

A second set of resolutions points to other frictions and behavioral assumptions. Coenen et al. (2017) note that the puzzle relies on perfect, "bounded rationality"; if agents are less forward-looking, the effects of future promises are naturally smaller. Similarly, Del Negro et al. (2012) propose that incorporating a "perpetual youth" (Blanchard & Yaari) structure, where agents have finite planning horizons and thus higher effective discount rates, generates "significantly more realistic responses" to future policy changes. (Del Negro et al. 2012)

A third resolution argues that the "puzzle" is a result of a false premise : it assumes the public only hears a pure Odyssean signal and therefore generates expansionary effects . In reality, as Andrade and Ferroni (2021) argue, any announcement is a mix of Odyssean commitment and Delphic information. They find that these two signals have opposite effects on macroeconomic expectations and financial conditions (Andrade & Ferroni, 2021). The puzzle, therefore, may not be a puzzle at all, but rather the result of standard models failing to account for the counterbalancing, contractionary "bad news" (Delphic) component that mutes the overall effect of the expansionary « Odyssean » signal.

2.2.3 Hybrid models

A more complex path of investigation that rejects the straightforward dichotomy has resulted from the realization that real-world forward guidance is never fully Delphic nor purely Odyssean. (Bassetto, 2019). According to this literature, the two forms are "intrinsically linked" and interact in intricate ways rather from being merely alternatives (Bassetto, 2019).

The author, for example, offers a critical game-theoretic perspective, arguing that in a world of symmetric information, meaning central banks have no superior information compared to the public, "pure Odyssean guidance is redundant" (Bassetto, 2019). A central bank, in this view, does not need to say it will commit; it can simply build credibility through its actions. Bassetto (2019) argues that forward guidance only becomes a valuable and necessary policy tool when the central bank has private information to convey specifically about its own preferences or policy beliefs. In this "hybrid" perspective, the central bank leverages its general credibility to make its Delphic announcements believable. This process is understood as "Odyssean guidance" when the specific information being disclosed is about the bank's own future intentions or commitment path.

This idea of an optimal mix is strongly supported by the welfare analysis of Goy et al. (2018). Their simulations show that while pure Odyssean guidance is the most powerful tool for reducing the likelihood of a deflationary spiral (lowering the probability by 11 percentage points), it does so at the cost of increased macroeconomic volatility, which reduces social welfare (Goy et al. 2018). Purely Delphic guidance, while weaker (reducing deflation risk by only 9.5 points), is crucial for stabilizing and

coordinating expectations. They conclude that a mix of both Delphic and Odyssean elements is preferred in welfare-terms, as it balances the need for stimulus against the need for stability.

Furthermore, the literature suggests that Delphic guidance should not be dismissed as merely a "bad news" signal. Xu and Xu (2023), in their investigation of New Keynesian models, find that Delphic guidance (specifically, targeting inflation) can be highly effective and achieve lower welfare losses than conventional policy, if the forward horizon is kept short and the proportion of rational agents in the economy is high (Xu & Xu, 2023). In contrast, they find Odyssean guidance is more robust to the length of the forward horizon. This implies that the choice of guidance strategy is not just about type, but also about timing and the perceived nature of the economic audience.

2.3 Fed vs Riksbanks monetary policy's evolution

This rich and complex theoretical debate, however, is largely overshadowed by a more fundamental conflict: a profound "disconnect between the theory and practice of forward guidance" (Moessner et al., 2015). While theorists like Krugman (1998) , Eggertsson & Woodford (2003) and Woodford (2012) championed the powerful, commitment-based Odyssean framework as the optimal solution to the ELB/ZLB, central banks in practice actively resisted it. Fearing the massive loss of credibility that would result from breaking an explicit promise, practitioners valued flexibility and discretion above all else.

2.3.1 The divergent evolution of Riksbank and Fed communication

Although the theoretical literature, previously discussed, debates between Delphic forecasts versus Odyssean commitments, central banks in practice have not been afforded the privilege of adhering solely to theorical principles. Their communication strategies have been forged through crises, evolving in distinct and divergent paths. The Sveriges Riksbank and the U.S. Federal Reserve, the two subjects of this analysis, represent a near-perfect experiment in this divergence. The Riksbank has pursued a path of pure Delphic transparency, while the Fed has evolved episodically, adopting a strategically ambiguous and reactive form of communication. This section will look at how each bank's practical framework has changed over time, how their choices have affected the market, and the big contradictions that come up in the literature that looks at how well they work, all the way up to the start of the 2020 pandemic.

2.3.2 The Sveriges Riksbank communication

The Riksbank's modern communication framework is a direct product of its history, particularly the lessons in restoring confidence through clarity learned during its severe 1990s financial crisis (Andersson & Viotti, 1999). the Riksbank's post-independence financial stability framework is a direct legacy of the 1990s crisis. The transparency-based tools developed to manage that crisis were not temporary fixed. They became the institutional pattern for the Riksbank's modern approach to policy.As one of the first central banks to adopt a formal inflation-targeting mandate (Flug & Honohan, 2022), its 2007 decision to begin publishing its own "repo rate path" was a pioneering and logical extension of this transparency-first doctrine (Andersson & Hofmann, 2009; Moessner et al., 2014). This positions the Riksbank, alongside peers in New Zealand and Norway, as a principal supporter of « quantitative forward guidance », interpreted as a publication of an interest rate path. (Andersson & Hofmann, 2009).

The Riksbank's leadership has been consistently clear that this path is not an Odyssean commitment, but rather "a forecast, not a promise" (Andersson & Hofmann, 2009 ; Breman & Seim, 2025). This purely Delphic stance is presented as a "feature, not a bug" (Speech of Breman, 2025); the constant and public revision of the interest rate path is intended to be the ultimate proof of the bank's flexibility and responsiveness to new information, thereby building, rather than breaking, public trust (Flug & Honohan, 2022). This flexibility is visually captured in Appendix 2. The "spray" of dotted forecast lines (black and grey) constantly deviating from the actual policy rate (blue line) is the signature of a Delphic framework in action. It is a visual admission of uncertainty and a rejection of strict commitment (Flug & Honohan, 2022). The Riksbank's transparency model, which also includes publishing detailed, attributed minutes of policy meetings and alternative scenarios has led the Riksbank to be "often ranked highly in measurements of monetary policy transparency" (Breman & Seim, 2025).

However, a critical examination of the literature reveals an unresolved debate regarding the actual effectiveness of this Delphic tool. The literature contradicts itself, calling into question the real-world utility of the Riksbank's core strategy. A body of research supports the Riksbank's view. Andersson and Hofmann (2009), in an early assessment, suggest that publishing a rate path "may enhance the central bank's leverage over medium-term interest rates"(Andersson & Hofmann, 2009). Moreover, Moessner and al. (2014) conclude that financial markets correctly understand the Riksbank's Delphic intention. Their analysis found that the publication of their path did not significantly affect the sensitivity of interest rates to macroeconomic news (Moessner et al., 2014). This implies markets wisely continued to react to data, not to the path itself, viewing the path as the conditional forecast it was intended to be (Moessner et al., 2014).

Yet, a more critical and arguably more damning set of analyses directly challenges these conclusions, suggesting the repo rate path may be an ineffective, or even counterproductive tool. Natvik et al. (2019), in a rigorous decomposition of market reactions in both Norway and Sweden, find no evidence that publishing interest rate projections improves markets' ability to forecast future rates. Their methodology, which separates market reactions into a "target" factor (response to the current rate decision) and a "path" factor (response to the future guidance), concludes that any forecast improvements stem from the current action, not the future path.(Natvik et al., 2019). This finding questions the Riksbank's stated objective for this policy, suggesting that it may be a transparent but informationally meaningless exercise.

This critique is further substantiated by Woodford (2012), who noted that « the Riksbank's time-contingent guidance was often unsuccessful, as market expectations frequently failed to align with the projected policy rates » (Woodford, 2012 ; cited in Moessner et al., 2014). Goodhart & Rochet (2011) went even further, suggesting that longer-term market rates do not react to surprises in the official path, indicating that "the projected repo path at longer horizons adjusts to market rates, rather than the other way round" (Goodhart & Rochet, 2011; cited in Moessner et al., 2014). This turns the entire logic of forward guidance on its head: instead of the Riksbank guiding the market, the market may be guiding the Riksbank's forecasts.

The fragility of this Delphic credibility was put on dramatic display during the 2011-2013 period, an episode detailed by Svensson (2015) as a "great failure" of Riksbank communication. Appendix 3 provides the definitive evidence. The Riksbank, pursuing its controversial "leaning against the wind" policy to manage household debt, published a repo rate path (solid black line) projecting a steep tightening. The market (dashed black line), viewing this as fundamentally non-credible given the weak inflation and employment outlook, completely ignored the guidance and correctly priced in a steep decline in rates. The Riksbank's justification for its path was, in itself, based on an incorrect forecast for foreign policy rates (grey lines), which were also far above market expectations (Svensson, 2015). The Riksbank was ultimately forced to abandon its own forecast and follow the market's trajectory, crashing its guidance credibility (Svensson, 2015). This episode reveals the critical weakness in the

Riksbank's Delphic model: its effectiveness is not absolute but is conditional on the market's belief in the credibility of its economic assumptions. When those assumptions are seen as inaccurate, the guidance, no matter how transparent it is , becomes irrelevant.

2.3.3 The U.S. Federal Reserve communication

In contrast to the transparency displayed by the Riksbank, the communication strategy of the US Federal Reserve has been marked by crises and occasional changes. Its journey has been from a pre-1994 culture of near-total secrecy with no communication (Campbell et al., 2017 ; Masciandaro et al.,2024) to a post-2008 framework defined by a multitude of unconventional tools and ambiguous signals. Unlike the Riksbank's interest rate path, the Fed's approach to communication has been multifaceted, uncoordinated, and, as a result, exceptionally difficult to interpret.

The 2008 global financial crisis and the descent toward the effective lower bound were the main triggers, forcing the Fed to adopt two important tools ,which were forward guidance and large-scale asset purchases (LSAPs) , as its main policy levers. (Boards of governors of the Federal Reserve System; Labonte, 2021). The Fed's forward guidance during this period, was a case study in reactive policymaking. It evolved through three distinct, experimental phases, classified by Moessner et al. (2015) as « Aesopian forward guidance »:

- Open-ended forward guidance : From December 2008, the Fed offered a vague, qualitative promise to keep rates low "for some time." Overall, this type of guidance is ambiguous and doesn’t provide detailed information about the time frame. (Moessner et al., 2015)

- Calendar-based/ Time-contingent forward guidance : in August 2011,vague promise shifted to a more concrete, time-contingent promise to hold rates low "at least through mid-2013" (a date later extended to "mid-2015"). This, however, was criticized for being disconnected from market participants’ expectations (Woodford, 2012).

- Threshold-based/ State-contingent forward guidance : In December 2012, it pivoted again to the "Evans rule," a data-dependent threshold, promising to keep rates at zero "at least as long as the unemployment rate remains above 6.5 percent" and inflation projections remained contained. This approach allows the public to see whether a change in policy is a response to a new economic outlook or a shift in the central bank's own preferences,a distinction that the two other guidance cannot clarify. (Moessner et al., 2015).

This progressive approach stands in direct opposition to the Riksbank's unified framework. The problem is exemplified by the Fed's main quantitative tool, the Summary of Economic Projections (SEP) which includes all FOMC participants’ economic projections, in which « the dot plot » is published.(Federal Reserve System Publication). This « dot plot » indicates how likely the monetary policy is going to evolve in the future according to the Fed (Bernanke, 2025). As the June 2019 dot plot (Appendix 4) illustrates, this is not a unified committee forecast, but an anonymous, "confusing spread" of individual policymakers’ projections (Svensson, 2015 ; Bernanke, 2025). It presents inconsistent projections derived from a contradictory set of individual opinions, which may differ from the chair's statement on monetary policy during press conferences (Bernanke, 2025). This lack of unified forecasts creates a significant “communication gap” when compared to other peers (Bernanke, 2025).

However, this ambiguity does not mean that the Fed's communication is without power. On the opposite, the literature shows that its impact on the markets is considerable, although it is extremely

complex to isolate. Swanson (2021), in a comprehensive study from 1991-2019, finds that both forward guidance and LSAPs had "substantial, highly statistically significant and persistent effects on a wide range of asset prices », effectively serving as substitutes for conventional policy. Neely (2014) confirms this impact, finding large international effects with Fed unconventional policy announcements, significantly lowering international bond yields and depreciating the U.S. dollar (Neely, 2014).

The Fed's guidance therefore works but how it works is a subject of intense critical debate.The problem, as identified in the literature, is identifying the channel. The Fed's impact is confounded by two types of "noise" that are absent from the Riksbank's framework:

The "Fed Information Effect," identified by Nakamura and Steinsson (2013), presents a powerful critique. Using a high-frequency analysis of market reactions, they find a paradox: a surprise Fed's increase in interest rates paradoxically causes markets to raise their output growth forecasts. They argue that markets are not just reacting to a "policy signal", which is the raise here, but also to an "information signal." The public interprets that as the Fed's private, optimistic assessment of the economy's underlying strength. This effect fundamentally complicates the interpretation of any Fed announcement and suggests that attempts to measure the pure "policy shock" are likely misspecified, as they are contaminated by this "information effect."(Nakamura & Steinsson, 2013)

The "Many Voices" Problem: Unlike the Riksbank's single path, the Fed's communication is unclear. Swanson & Jayawickrema (2024) make a contribution by demonstrating that official FOMC announcements are not the main source of policy shocks. Their analysis finds that for longer-term assets (which are the target of forward guidance), speeches by the Fed Chair and succeeding press conferences often move markets as much or more than the announcements themselves. This highlights the fact that an analysis based solely on FOMC statements is incomplete.

Consequently, the Fed's pre-pandemic framework was characterized by ambiguity,. Its guidance was episodic, only used during crises to handle ZLB/ELB situations. its quantitative signal (the dot plot) was a contradictory set of individual opinions, and its impact on the market, while strong, was inseparably linked to information effects, the dispersion of communication sources, and the inherent limitations of the credibility of its own imperfect communication. This framework was about to collide with one of the deepest economic shock of the century.

2.4 The Covid-19 shock

2.4.1 The shock and diverse policy responses

The Covid-19 pandemic was an economic cataclysm that triggered levels of uncertainty and volatility unseen in modern financial history.Both the Riksbank and the Fed responded with huge measures. The Fed's communication, in particular, became more reactive, focused, and synchronized with its unconventional policy actions than in previous crises (Benchimol et al., 2025).

Yet, in the field of communication strategy, their responses were different, providing an opportunity to analyze their contrasting philosophies. The Riksbank responded by strengthening its established Delphic framework. Its actions were seen as effective and decisive in mitigating the crisis (Flug & Honohan, 2022). Critically, it used the crisis as an opportunity to enhance its Delphic transparency, not abandon it. As advocated by its own leadership (Breman, 2025) the Riksbank made the publication of alternative scenarios a standard feature of its Monetary Policy Reports since 2023. This is opposite to

the vision of an Odyssean commitment. It is an explicit acknowledgment of deep uncertainty, a move to communicate how policy might change, rather than a rigid promise that it won't.

The Federal Reserve, on the other hand, did exactly the opposite. It entered the pandemic after adopting a new revised framework in August 2020, the “Statement on Longer-Run Goals and Monetary Policy Strategy” (Board of Governors of the Federal Reserve System, 2021). This revised framework, which introduced a flexible average inflation target (FAIT) aimed at stabilizing inflation at 2% and also emphasized “shortfalls” rather than “ deviations” from maximum employment, was the intellectual product of the previous crisis (English & Sack, 2024).
As English and Sack (2024) and Romer and Romer (2024) argue, it was a framework "too focused on the experience following the financial crisis" (English & Sack, 2024), designed to tackle chronic disinflation following the last crisis. It proved to be catastrophically "not robust" when confronted with unexpected changes such as the high-inflation shock of the pandemic (English & Sack, 2024) and therefore, performed poorly (Romer & Romer, 2024).

The Fed's implementation of its 2020 framework constituted a failed experiment that can be interpreted as an odyssean commitment ,with some promises.As detailed by Labonte (2021) and Papell and Prodan (2024), the Fed issued explicit, “outcome-based” forward guidance: it publicly committed to hold rates at the zero-lower bound/effective lower bound until two conditions were achieved: first, its ambitious "maximum employment" goal, and second, inflation had risen to 2% and was "on track to moderately exceed 2% for some time." (Labonte, 2021). These declarations were done, regardless of the circumstances.

This was, by definition, a state-contingent forward guidance (seen above). The Fed was publicly "tying itself to the mast”, committing to let the economy "run hot" and explicitly not act before in order to achieve its new FAIT goal (Romer & Romer, 2024; Federal reserve Board, 2025). As inflation surged throughout 2021, the Fed remained notoriously "behind the curve" (Papell & Prodan, 2024). The critical analysis of Romer and Romer (2024) argues that the Fed was constrained by its own guidance. The "elevation of the maximum employment goal" and the explicit move "away from preemptive policy actions " left the FOMC trapped, forced to wait for its employment outcome to be achieved while inflation spiraled out of control.(Romer & Romer, 2024)

The market's reaction provides a damning verdict on this Odyssean failure. As documented by Bauer et al. (2024), markets and professional forecasters did not update their beliefs about the Fed's hawkishness in response to its communication in late 2021. The Fed's "talk" was not credible. Perceptions of its policy rule only shifted after the Fed was forced to violently abandon its guidance and "liftoff" with its first-rate hike in March 2022. This "behind the curve, pivot, and getting back on track" pattern (Papell & Prodan, 2024) had a devastating effect on market stability. Appendix 5, from Bauer et al. (2024), provides the stark evidence. The chart shows that after a period of relatively contained policy surprises, the market's "surprise" factor exploded in 2022 as the Fed was forced to pivot. The guidance, intended to create stability and "tie the bank to the mast," had instead done the exact opposite.

2.4.2 Research Gap and Sentiment analysis

The pandemic and the Fed's failed experiment have exposed the deep, practical limits of forward guidance. The literature suggests that in times of extreme uncertainty, the tool itself may break down. Levin et Sinha (2020) argue that during the Covid-19 crisis, forecaster "myopia" and disagreement were so high that markets already expected short-term rates to remain at the ELB for several years (Levin & Sinha, 2020). In such an environment, Delphic guidance has no additional effect. To have any impact,

the bank would need to promise extreme (and thus non-credible) future overshooting of its targets, highlighting a severe limitation to Odyssean policy. The Fed's 2020 guidance demonstrated the opposite, but fatal, weakness ; as Labonte (2021) notes, when guidance is too specific and outcome-based, it loses credibility at the moment the economic situation changes, imposing a policy shift that destroys market expectations rather than consolidating them (Labonte, 2021).

English and Sack (2024) call for a simpler, "constitutional" framework and an end to outcome-based guidance(English & Sack, 2024). Romer and Romer (2024) support this idea, stating that explicit guidance on future policy could be entirely “unnecessary.” They suggest that the Fed should simply clarify its objectives and forward guidance, then let markets deduce its path, using tools such as the SEP and speeches to provide additional information, rather than as binding constraints (Romer & Romer, 2024).

In a major critique, Bernanke argues the Fed should abandon the confusing "dot plot" and instead publish a single, unified, staff-led "Economic Review," the same principle used by the Riksbank . Both Bernanke and Mester (2024) strongly advocate for the Fed to adopt the use of alternative scenarios (Bernanke, 2025; Mester, 2024). Again, it is a tool the Riksbank already actively integrates (Breman, 2025). This represents a clear intellectual concession regarding the failure of the Fed's episodic model and the fact that a scenario-based approach may be more robust in an uncertain world.

However, a critical gap remains. The literature lacks a quantitative measure of the communication itself, and has not yet provided a concrete method to measure the degree of "Odysseanness". Numerous studies have shown that FG is solely binary, whereas it is important to verify that it is instead a hybrid approach .This thesis will try to address this gap, empirically answering the research question: "has forward guidance become more Odyssean?". To do so, I will use sentiment analysis through sentiment index’s scores derived from an analysis of the word choice and linguistic patterns used in official speeches. This methodology will be applied in a comparative study of the Federal Reserve and the Riksbank, two major central banks in terms of forward guidance, whose approaches are fundamentally divergent. Crucially, the analysis will be structured as a pre/post- Covid-19 comparison, treating the pandemic as a natural experiment. This will allow us to quantitatively measure how communication in these two distinct frameworks has evolved and how it has presented up to now.

3. Methodology

3.1 Methods

3.1.1 Research Design

Consistent with a deductive research approach, this study employs an explanatory quantitative design to empirically test theoretical hypotheses regarding the impact of central bank communication on financial markets. The research framework is operationalised through sentiment analysis, which serves to quantify the informational content of speeches by constructing sentiment score indices. Within the sentiment analysis family, the study adopts a lexicon-based approach (Mao et al., 2024) which has often been used in the established literature and its ability to provide a transparent, replicable procedure that maintains scientific rigour. These sentiment indices subsequently serve as the primary explanatory variables in an Ordinary Least Squares (OLS) regression framework, allowing for a deeper study of whether variations in communication tone are associated with changes in government bond yield volatility.

Regarding the scope of the study, while the initial objective was to conduct a symmetric comparative analysis of both banks, the raw data revealed a significative asymmetry in speech volume: the initial Federal Reserve corpus contained 765 speeches, whereas the Sveriges Riksbank provided 109. As a consequence , the empirical analysis places its primary focus on the Federal Reserve, using the Riksbank as a complementary case study to explore institutional or structural differences in a low-power setting. Following a rigorous filtering process designed to isolate communications specifically addressing forward guidance, the final economic research sample consists of approximately 194 speeches for the Federal Reserve and 38 for the Riksbank.

3.2 Data Selection

The first part of this analysis focuses on the treatment of all speeches delivered by the FED and the Riksbank, with an emphasis on the period 2012-2024 and will then be restricted later on. The choice of this period is strategically determined by the need for a consistent and theoretically harmonized environment for both banks. As previously seen in the literature, this period corresponds to the “explicit era” of forward guidance, established by the introduction of the Odyssean-Delphic dichotomy by Campbell et al. (2012). For the Federal Reserve, 2012 coincides with the first official publication of the "dot plot “ (Svensson, 2015), which added a quantitative dimension to its communication. For the Riksbank, this year marks the end of the highly controversial episode of ”leaning against the wind" (2010-2011), which reflects the bank's efforts to refine its Delphic identity. (Svensson, 2015). Empirically, this 12-year horizon provides the structural variation needed to study the tone of communication in three distinct macroeconomic regimes common to both banks: the post-crisis zero interest rate environment, which includes the Riksbank's experience with negative rates (starting in 2015), followed by the Covid-19 period and the 2022-2024 inflationary cycle that followed. Finally, the early 2012 start date ensures methodological consistency: prior to that, the uncertain environment and changing communication norms at both banks introduced significant noise, further compromising the effectiveness of dictionary-based indices.

3.2.1 Data collection

First, speeches needed to be available and collected within the 2 central banks' website.To do so, a method called "web-scraping", which consists of implementing codes to extract all the speeches directly , without doing that task by hand and save a lot of time, is chosen. On the one hand, The Federal Reserve maintains a well-structured speech archive in HTML format over a long horizon, which allows for straightforward and reliable automated extraction. By contrast, the Riksbank's archive is more complex, with speeches provided as PDF documents and material before 2017 stored in archived sections, which increases the risk of incomplete retrieval and format-related inconsistencies when scraping directly from the official website. To ensure systematic coverage and harmonised formatting for the historical component of the corpus, the study therefore uses the comprehensive central-bank speech repository compiled by Hansson (2021) as a data-acquisition process. This repository is employed strictly to retrieve the speech texts and associated data.To mitigate potential transcription or errors related to external sources, a validation protocol was applied: a random portion of speeches was compared with official central-bank sources (date, title, speaker) to ensure that nothing has been changed from the version on each bank's official website.

3.2.2 Data cleaning

To ensure the feasibility of the sentiment analysis, the raw text corpus was subjected to a rigorous multi-stage preprocessing pipeline implemented via Python. Prior to the cleaning procedures, a structured classification framework was established to organize the data for following empirical analysis. A bank's indicator was assigned to each observation (either FED or Riksbank), accompanied by two additional variables representing the speaker's identity and his institutional role. Specifically, a dichotomous classification was employed for the institutional role: the label "Head" was assigned to speeches delivered by the Chairs of the Federal Reserve (Bernanke, Yellen, or Powell) or the Governors of Sveriges Riksbank (Ingves or Thedéen). Conversely, the label "Member" was applied to all other officials within the respective institutions. This organizational stage was essential to facilitate later empirical testing of whether the speaker's role and institutional status influence the market's reaction to speech communication.

Following the methodology suggested by Siino et al. (2023), the first phase involved noise reduction, wherein automated codes[1] eliminated HTML links, typographical errors, and "polite formulas." Special characters were also removed to prevent the generation of non-lexical tokens.
A critical component of this preprocessing phase was contraction resolution, which specifically expands terms such as "shouldn't" to "should not." Within the context of monetary policy, this step is mandatory for improving accuracy; as central bank guidance often relies on negation, failing to isolate these terms could lead to a misinterpretation of policy intent. For instance, a sentence such as "we should not commit" would be wrongly classified as a positive "Odyssean" signal if the negation were not properly decoupled from the verb "commit."

Following steps included the removal of punctuation and numerical data, as these elements do not contribute to the qualitative evaluation of commitment or forecasting. The entire corpus was then converted to lowercase to ensure that the algorithm treated identical terms, such as "Commitment" and "commitment", as a single observation. Interestingly, this study deviated from standard Natural Language Processing (NLP) techniques by retaining certain "stop words" and avoiding lemmatization. While stop words are typically removed because they represent the majority of words in a text, modal auxiliaries (e.g., "should," "will") and temporal prepositions (e.g., "until," "during") were preserved here, as they carry significant weight in establishing the central bank's tone and degree of certainty.

[1] All the code is provided in a supplementary file

Similarly, lemmatization and stemming, which consist of cutting words to their morphological roots, were avoided to maintain compatibility with the chosen sentiment lexicon, which is not structured in a lemmatized format. Applying these techniques would have resulted in unrecognized "cut" words, thereby reducing the precision of the dictionary-based matching (Hardeniya & Borikar, 2016).

The final phase of preprocessing involved tokenization, implemented through the Python Natural Language Toolkit (NLTK) (Vijayarani & Janani, 2016). While the initial design explored the use of "n-grams" to capture multi-word expressions like "Federal Reserve System," the study primarily utilized "unigrams" (single-word tokens) to minimize the "noise" generated by improperly linked phrases. To compensate for this transition and preserve the precise meaning of key monetary policy concepts, a custom "protected list" was developed[2]. This allowed the algorithm to treat essential compound terms such as "forward guidance" and "zero lower bound" as single, intact units of meaning, ensuring that the structural nuances of monetary speeches were preserved throughout the analysis.

3.2.3 Sentiment analysis

After the data were cleaned and tokenized, sentiment indices were constructed using a lexicon-based approach to sentiment analysis. This family of methods comprises a dictionary-based and a corpus-based approach.The former was selected in this thesis because of its feasibility and its widespread use in the empirical literature (Mao et al., 2024). In line with the general objective of sentiment analysis, which consists of automatically extracting and classifying opinions and sentiments from text using Natural Language Processing techniques (Hardeniya & Borikar, 2016), the aim here is to classify tokens into categories that capture the tone of central bank communication and to compute aggregate indices at the speech level. While existing work often focuses on broad dimensions such as positive versus negative, or hawkish versus dovish tone, this thesis requires a more subtle distinction between Odyssean guidance (language reflecting commitment to a future policy path) and Delphic guidance (language relecting forecasts and uncertainty about the economic outlook).

Because no pre-existing dictionary directly measures Odyssean and Delphic tone, it was necessary to adapt and extend a finance-specific lexicon. The Loughran and McDonald (LM) dictionary offers a suitable basis, as it provides domain-specific lists for negative, positive, uncertainty, litigious, and strong/weak modal words, calibrated on financial and corporate contents rather than general language. This lexicon corrects the systematic biases of generic dictionaries and instead isolates words that are genuinely associated with bad news, uncertainty, legal risk, or modal strength in financial contexts (Loughran & McDonald, 2011). Its uncertainty category is naturally related to Delphic communication, while its strong and weak modal categories are particularly relevant for capturing degrees of commitment characteristic of Odyssean guidance. However, LM alone does not resolve the core problem of distinguishing language that conveys commitment from language that conveys forecasts and uncertainty in central bank speeches.

To address this, the methodology draws on Hansen and McMahon (2016), who use the LM dictionary as a foundation and, in parallel, construct a forward-guidance index applied only to specifically identified guidance sections, thereby ensuring that LM-type categories are interpreted within a forward-looking policy context. Other contributions, such as Benchimol et al. (2025), Hubert and Fabien (2017), Correa et al. (2017), Apel and Grimaldi (2012), and Bernard et al. (2022), similarly rely on or adapt LM-style financial dictionaries for central bank communication, reinforcing the suitability of this approach for the current thesis. Building on this literature, a custom dictionary was developed to capture Odyssean and Delphic tone, and then integrated into the LM structure to compute approximate sentiment indices. The first version of the index measures the overall prevalence of Odyssean and Delphic language even outside explicit forward-guidance context, in order to assess how

[2] This list will be used as a filter, see later on (page 19)

the vocabulary associated with commitment and forecast-type communication has evolved over time across central bank speeches.

The construction of the custom dictionary proceeds as follows. Campbell et al. (2012), who introduced the Odyssean–Delphic distinction in the context of forward guidance, are used as the conceptual basis for compiling two lists of terms: one associated with commitment-oriented guidance (Odyssean) and one associated with forecast- and uncertainty-oriented communication (Delphic). These two lists are then merged into the LM dictionary as additional columns, alongside the existing uncertainty and modal categories. As a reminder, lemmatization was deliberately omitted at the cleaning stage because the LM dictionary already provides an extensive set of inflected forms and roots; aggressive lemmatization would have distort these forms and reduce matching accuracy. In the implementation, each tokenized speech is matched against the customized LM dictionary. Whenever a token appears in the Odyssean or Delphic lists, it is counted in the corresponding category. At the end of this process, each speech is associated with counts of Odyssean, Delphic, uncertain, negative, positive, and modal terms, from which normalized sentiment scores are computed. The *Table 1* represents a sample of the customized dictionary used for computing indices scores

Table 1 : Sample of the customized LM dictionary (2011).

Word	Negative	Positive	Uncertainty	Strong_modal	Weak_modal	Odyssean	Delphic
abandon	1	0	0	0	0	0	0
commit	0	0	0	0	0	1	0
doubt	1	0	1	0	0	0	1
empower	0	1	0	0	0	0	0
forecast	0	0	0	0	0	0	1
guarantee	0	0	0	0	0	1	0
might	0	0	1	0	1	0	1
must	0	0	0	1	0	1	0
possible	0	0	1	0	1	0	1
promising	0	0	0	0	0	1	0
scenario	0	0	0	0	0	0	1
variability	0	0	1	0	0	0	0
will	0	0	0	1	0	0	0
worthy	0	1	0	0	0	0	0

3.2.4 Sentiment Scores

After quantifying the corpus, sentiment indices are constructed at the speech level in order to track changes in the central bank's tone over the period 2012-2024. Rather than adopting a net balance index often used in this type of approach, such as $(Positive - Negative)/(Positive + Negative)$, this study differs from this practice for both conceptual and econometric reasons. Conceptually, Campbell et al. (2012) define Odyssean and Delphic signals as two distinct dimensions of communication, one expressing a commitment to a future policy stance and the other providing forecasts about the economic outlook, rather than as opposite ends of the same continuum. A given speech may therefore contain a high concentration of both Delphic and Odyssean language. A net

balance index would allow these signals to counterbalance each other, which could result in a neutral value even when both dimensions of orientation are strongly expressed.
To preserve this structure, 2 topic-proportion indices are defined by the following equations and inspired by Istrefi et al. (2023) :

$$OdyScore_d = \frac{\sum_{j=1}^{R_{Ody}} r_{Ody,d,j}}{W_d} \qquad (1)$$

$$DelScore_d = \frac{\sum_{j=1}^{R_{Del}} r_{Del,d,j}}{W_d} \qquad (2)$$

Where:

- d= a given speech within the analyzed corpus
- $OdyScore_d$= Odyssean score for speech $d \in [0; 1]$
- $DelScore_d$= Delphic score for speech $d \in [0; 1]$
- W_d = total number of words in speech d.
- R_{Ody}= number of Odyssean terms in the Odyssean lexicon
- R_{Del}= number of Delphic terms in the Delphic lexicon
- $r_{Ody,d,j}$= frequency of the j-th Odyssean term in speech d
- $r_{Del,d,j}$= frequency of the j-th Delphic term in speech d

Normalising by the total word count W_d ensures comparability across speeches of different length and preventing longer speeches from mechanically exhibiting higher tone indices simply because they contain more tokens. Treating Odyssean and Delphic scores as separate proportions is consistent with topic-proportion in central bank communication studies (Istrefi et al., 2023), where each dimension is modelled as an independent “topic” whose intensity varies over time. To ensure precision, a negation-handling heuristic was integrated into the scoring algorithm to prevent the misclassification of words based on the context. The script identifies specific markers, such as "not," "no," "never”, within a preceding three-token window of any Odyssean term. This mechanism ensures that phrases like "no commitment" are not erroneously recorded as positive odyssean signals. While this rule-based approach is not a perfectly reliable solution for capturing complex syntactic nuances, it serves as a transparent filter to mitigate signal
noise in the resulting sentiment indices. For the econometric analysis, both indices were further standardised within each central bank: Z-scores (Ody_Std, Del_Std) are obtained by subtracting the bank’s mean μ and dividing by the bank’s standard deviation σ. This transformation allows regression coefficients to be interpreted as the effect of a one-standard-deviation increase in Odyssean or Delphic tone on market volatility, and it improves the comparability of estimated effects across regressions, especially between the FED and the Sveriges Riksbank.

Once the Odyssean and Delphic scores had been calculated and standardized, the process was repeated for the uncertainty and negative scores. For future regressions, the inclusion of negative and uncertainty scores as independent variables is motivated both by theory and by existing evidence on how central bank communication affects markets. Negative tone is likely to be correlated with market volatility and reactions in similar contexts. The tone of uncertainty , meanwhile, reflects the extent to which the central bank emphasizes ambiguity regarding the outlook or monetary policy, a dimension that is particularly relevant to volatility. Loughran and Mcdonald (2012), established a solid foundation showing the real impact of negative and uncertain words on market volatility. The use of these two variables in regressions therefore allows the analysis to verify if episodes of high Odyssean or Delphic guidance coincide with more negative or uncertain communication, and to distinguish the specific effect of the tone of guidance from the general sentiment context.

All previous preprocessing steps were designed to support the quantitative analysis of how forward guidance tone affects financial markets. While it is interesting to document the general evolution of Odyssean and Delphic tone across all speeches, the primary objective of this thesis is to analyse how explicit FG communication operates in specific macroeconomic environments, particularly when policy rates are constrained by the ELB/ZLB , and how such guidance influences market reactions. In the literature above, FG was treated as an unconventional policy instrument that was particularly important at the ELB/ZLB, but also complements or partially substitutes for conventional policy depending on the economic context. For this reason, the empirical analysis narrows the focus to speeches that explicitly address forward guidance, rather than all monetary policy speeches that merely contain strong commitment or forecast language. To implement this restriction, a dedicated FG filter was built consisting of 15 carefully selected forward-guidance–related terms; *Table 2* reports the list of these key words, which was deliberately kept narrow to avoid capturing speeches whose content is only loosely related to forward guidance. This filter was then applied to the corpus and the final sample is finally composed of 194 FG speeches for the Federal Reserve and 38 for the Riksbank.As previously mentioned , the riksbank sample is very small to support strong conclusions and will therefore be treated as a comparative case study within the regression analysis.

Table 2 : FG filter keywords

ELB/ZLB context	Unconventional Policy tool	Implementation and Communication
lower_bound	forward_guidance	liftoff
effective_lower_bound	quantitative_easing	balance_sheet_normalization
zero_lower_bound	large_scale_asset_purchases	federal_funds_rate
yield_curve_control	asset_purchase_program	repo_rate_path
	asset_purchase_programme	monetary_policy_rules
		dot_plot

3.3 Variable Selection

3.3.1 dependent variables

The empirical analysis measures market reactions using daily movements in medium-term government bond yields. For the United States, daily 2-year and 5-year Treasury constant-maturity yields are obtained from the Federal Reserve Economic Data (FRED) database (series DGS2 and DGS5). For Sweden, daily 2-year and 5-year Swedish government bond yields are collected from official yield series published through the Riksbank's statistical sources. The 2-year maturity is treated as the primary outcome because forward guidance is expected to transmit mainly through expectations about the near-term policy rate path, which is reflected most directly in short- to medium-term yields (Gürkaynak, Sack, and Swanson, 2005; Swanson, 2017). The 5-year maturity is included as a second maturity comparison only.[3]

Daily yield changes are computed as $\Delta Y_t^{(m)} = Y_t^{(m)} - Y_{t-1}^{(m)}$, where $t-1$ denotes the previous available trading day and $m, the\ maturity$. To capture the volatility of yield move regardless of direction, the dependent variables are defined as absolute changes $|\Delta Y_t^{(m)}|$ (Cieslak & Schrimpf, 2019). Observations for which $\Delta Y_t^{(m)}$cannot be computed due to missing yields (e.g., weekends, holidays,..) are removed when constructing the regression dataset. Finally, speech dates are mapped to the closest relevant market close: when a speech is treated as arriving after the market close (or on a non-trading day), the reaction is shifted to the next trading day to avoid attributing to the speech a yield movement that occurred before markets could plausibly process the information.

3.3.2 Explanatory variables and controls

The econometric analysis relies on four standardized tone measures computed at the speech-day level: Odyssean tone (Ody_Std), Delphic tone (Del_Std), negative tone (Neg_Std), and uncertainty tone (Unc_Std). The two forward-guidance measures, Ody_Std and Del_Std, are the main explanatory variables, capturing respectively commitment-oriented and forecast-oriented language in FG speeches. Neg_Std and Unc_Std are included as control variables to hold sentiment dimensions that are commonly related to yield volatility, so that the estimated associations for Ody_Std and Del_Std are not simply picking up general negativity or uncertainty in the communication (Loughran & McDonald, 2012).

Because yield outcomes are measured at daily frequency, the unit of observation is a speech-day for each central bank. When several FG speeches occur on the same calendar day within the same institution, they are aggregated into a single observation to ensure a one-to-one match between that day's communication and the corresponding daily yield change, and to avoid duplicating the same yield movement across multiple speeches. Consistent with this rule, the speaker-status indicator is defined at the day level as $\mathrm{Head}_d = \max_{s \in d}(\mathrm{Head}_s)$, so that $\mathrm{Head}_d = 1$ whenever at least one FG speech on day d is delivered by the Chair/Governor depending on the bank, and $\mathrm{Head}_d = 0$ otherwise. Two additional controls are also used. First, a Covid-19 dummy variable is defined as $\mathrm{Covid}_d = 1$ for dates from 1 March 2020 to 31 December 2021 (included) and 0 otherwise, to capture the major structural break and unusually high market volatility during the pandemic period. Second, because speech events are not evenly spaced over time, the analysis controls for the number of days since the previous speech-day within the same central bank sample, with the variable $\mathrm{GapDays}_i = d_i - d_{i-1}$, and uses

[3] It avoids confusion as larger maturities may be influenced by other risk and growth factors

the transformation $\log(\text{GapDays}_i)$ to reduce right-skewness and limit the influence of very long gaps (Wooldridge, 2009).

3.3.3 Robustness checks

To assess whether the regression results are sensitive to the exact definition of forward-guidance tone, a structured set of robustness checks is implemented for the baseline specification. In applied econometrics, robustness checks serve as diagnostic tools to evaluate how core coefficients behave when tone measurement is modified, without claiming structural proof of validity (Lu & White, 2013)

- **Tone-specification robustness**. The baseline specification is re-estimated with alternative tone variable sets: 1) excluding Uncertainty tone, 2) excluding Delphic tone, and 3) replacing Odyssean and Delphic tone with a single relative index NetFG = Ody_Std − Del_Std. These alternatives verify that the main results are not mechanically driven by correlation between forecast-type and uncertainty-related language, and whether effects depend on treating the two tones jointly or through their difference. Such sensitivity checks are consistent with best practices in applied econometrics, aiming for evaluating coefficient stability (Lu & White, 2013).

3.4 Regressions and Hypotheses

This section develops testable hypotheses derived from the literature review and specifies the regression models used to estimate the association between forward-guidance tone and daily absolute yield changes. The empirical strategy adopts an event-style regression approach, consistent with Picault and Renault (2017), and is adapted to speech-based sentiment indices.

3.4.1 Theoretical Hypotheses [4]

Before specifying the estimation equations, we derive hypotheses regarding the relationship between communication tone and the magnitude of market reactions.

- **H1 and H2: Tone and absolute yield changes.**
  Predictions for the sign of the association between tone and absolute yield changes $(|\Delta Y|)$ are ambiguous. Credible Odyssean commitments may reduce policy uncertainty and dampen yield movements (Campbell et al., 2012) but strong commitments may also represent "information effect" that triggers immediate repricing (Nakamura & Steinsson, 2013). Likewise, Delphic guidance can increase repricing by revealing information, or stabilize expectations by clarifying the reaction function (Goy et al., 2018). Therefore, we test non-directional hypotheses: $H1$: Odyssean tone is significantly associated with the magnitude of daily yield changes $(\beta_1 \neq 0)$. $H2$: Delphic tone is significantly associated with the magnitude of daily yield changes $(\beta_2 \neq 0)$.

- **H3 and H4: Control variables (negative and uncertainty tone).**
  In contrast, the literature suggests that expressions of negativity and explicit uncertainty increase market anxiety (Loughran & McDonald, 2012). We therefore expect these dimensions to be positively associated with the magnitude of yield movements:

[4] Note : *Hypotheses H1–H7 are economic hypotheses based on the literature. They will be tested later on to assess their significance*

$H3$: Negative tone is expected to be positively associated with absolute yield changes ($\beta_3 > 0$).
$H4$: Uncertainty tone is expected to be positively associated with absolute yield changes ($\beta_4 > 0$).

- **H5 and H6: Speaker role**
  Swanson and Jayawickrema (2024) show that Fed Chair speeches carry greater weight in financial markets than those of other FOMC members. We therefore expect speaker authority to alter the association between tone and absolute yield changes. In the interaction model (F2), this implies that the marginal association between tone and yield changes differs when the speaker is the Chair:
  $H5$: Speaker authority alters the association between Odyssean tone and absolute yield changes ($\theta_1 \neq 0$), and the absolute marginal association $|\beta_1 + \theta_1|$ is expected to be larger for Chair speeches.
  $H6$: Speaker authority significantly alters the association between Delphic tone and absolute yield changes ($\theta_2 \neq 0$), and the absolute marginal association $|\beta_2 + \theta_2|$ is expected to be larger for Chair speeches.

- **H7: Covid regime**:
  The Covid-19 pandemic represented a period of extreme uncertainty and a distinct monetary policy regime. Motivated by evidence of altered policy transmission during this period (Romer and Romer, 2024; Bauer et al., 2024), we test whether the market sensitivity to Odyssean and Delphic signals differs during the crisis window:
  H7: The association between tone and absolute yield changes differs significantly during the Covid-19 period compared to the pre-pandemic baseline ($\varphi_1 \neq 0 \; and \; \varphi_2 \neq 0$).

3.4.2 Baseline Model

For each central bank, the baseline specification relates the absolute yield change on speech-day d and maturity $m \in \{2,5\}$ to the standardized tone indices and controls:

$$\left|\Delta Y_d^{(m)}\right| = \alpha + \beta_1 Ody\_Std_d + \beta_2 Del\_Std_d + \beta_3 Neg\_Std_d + \beta_4 Unc\_Std_d + \beta_5 Head_d + \beta_6 \log(GapDays_d) + \beta_7 Covid_d + \varepsilon_d \quad (3)$$

The Federal Reserve baseline is estimated with m = 2 and the same specification is also estimated with m = 5 as a maturity comparison to assess whether tone associations differ across the yield curve. For the Riksbank, the same baseline structure is estimated for m = 2 and m = 5 . Given the smaller Swedish sample, the Riksbank regressions are interpreted as an exploratory comparative benchmark rather than as a fully symmetric test. Because tone measures are standardized, β_1 and β_2 capture the association between a one-standard-deviation difference in Odyssean and Delphic tone and the magnitude of daily yield changes, conditional on controls. The dependent variable is the absolute daily change, which serves as a proxy for daily yield volatility.

3.4.3 Heterogeneity extensions

To examine whether the tone–yield relationship varies with speaker authority $(H5, H6)$, the baseline model is augmented with interactions between tone and the Chair indicator (model F2):

$$\left|\Delta Y_d^{(2)}\right| = \alpha + \beta_1 Ody\_Std_d + \beta_2 Del\_Std_d + \beta_3 Neg\_Std_d + \beta_4 Unc\_Std_d + \beta_5 Head_d + \beta_6 \log(GapDays_d) + \beta_7 Covid_d + \theta_1(Ody\_Std_d \times Head_d) + \theta_2(Del\_Std_d \times Head_d) + \varepsilon_d \quad (4)$$

Where the coefficients $\theta_1 \; and \; \theta_2$ capture how the association between tone and absolute yield changes differs when the speaker is the Federal Reserve Chair. To test for regime dependence during the Covid -19 period ($H7$), the baseline model is augmented with interactions between tone and the Covid indicator (model F3):

$$\left|\Delta Y_d^{(2)}\right| = \alpha + \beta_1 Ody\_Std_d + \beta_2 Del\_Std_d + \beta_3 Neg\_Std_d + \beta_4 Unc\_Std_d + \beta_5 Head_d + \beta_6 \log(GapDays_d) + \beta_7 Covid_d + \varphi_1(Ody\_Std_d \times Covid_d) + \varphi_2(Del\_Std_d \times Covid_d) + \varepsilon_d \quad (5)$$

Where the coefficients $\varphi_1 \; and \; \varphi_2$ measure how the association between tone and yield changes differs during 2020–2021 relative to the non- Covid baseline.

3.4.4 Alternative specifications

Beyond the maturity variants presented above, two alternative specifications assess the sensitivity of the baseline results. First, for the Fed only, a next-trading-day outcome is used to account for potential delayed market processing :

$$\left|\Delta Y_{d+1}^{(2)}\right| = \alpha + \beta_1 Ody\_Std_d + \beta_2 Del\_Std_d + \beta_3 Neg\_Std_d + \beta_4 Unc\_Std_d + \beta_5 Head_d + \beta_6 \log(GapDays_d) + \beta_7 Covid_d + \varepsilon_{d+1} \quad (6)$$

This variant addresses the possibility that some speeches occur after market close or that the market's response is delayed until the next session, given that intraday timestamps are unavailable to be fully precise

Second, tone-specification robustness is assessed by estimating alternative versions of the baseline model F1 that (i) exclude Unc_Std, (ii) exclude Del_Std, and (iii) replace Ody_Std and Del_Std with a single relative index $NetFG\_d = Ody\_Std\_d - Del\_Std\_d$. These alternatives verify that the main results are not mechanically driven by correlation between tone indices or by treating Odyssean and Delphic dimensions jointly rather than through their differences.

3.4.5 Estimation and inference

All models are estimated by OLS and statistical inference is based on heteroskedasticity-robust standard errors using the HC3 correction (MacKinnon & White, 1985; Hayes & Cai, 2007). Compared with the classical White (1980) correction, HC3 provides more reliable inference in small or moderate samples by adjusting residuals for observation leverage, thereby reducing the risk of overstated significance (Hayes & Cai, 2007). This correction is particularly appropriate given the unequal sample sizes between the Fed and the Riksbank and the presence of high-volatility outliers around major policy events, ensuring robust and conservative inference under heteroskedasticity across regression models.

Multicollinearity among regressors is assessed using variance inflation factors (VIFs), which quantify the extent to which linear dependence inflates the variances of the estimated coefficients (Wooldridge, 2009). Individual hypotheses on regression parameters are evaluated using two-tailed Wald tests of the null hypothesis $H_0: \beta = 0 \; against \; H_1: \beta \neq 0$; directional expectations are assessed using the estimated sign alongside this conservative two-tailed decision rule to maintain a single inference standard across coefficients.

For combined hypotheses involving multiple coefficients, heteroskedasticity-robust Wald tests are reported. Statistical significance is assessed at the 10%, 5%, and 1% levels. Regression results are presented in Chapter 4.

3.4.6 Marginal effects from interaction models

For interaction specifications presented in section 3.4.3, marginal effects were computed to evaluate the tone–yield association within each regime separately. In the reference regime (Head = 0 or pre-Covid), the marginal effect equals the main tone coefficient ($\beta_1 \, or \, \beta_2$). In the alternative regime (Head = 1 or during Covid), the marginal effect equals the main coefficient plus the interaction coefficient ($\beta_1 + \theta_1 \, or \, \beta_2 + \theta_2$).Standard errors for these linear combinations were obtained from the variance-covariance matrix of the HC3 robust coefficient estimates, accounting for the covariance between main effects and interaction terms.

4. Empirical Results

4.1 Descriptive statistics and tone dynamics

This section presents the empirical evidence on how forward-guidance communication has evolved over time and how it relates to the magnitude of government-bond yield reactions. The objective is to describe the findings using graphs and tables consistent with the research question and verifications of the findings but without yet linking them to the existing literature.

4.1.1 Descriptive statistics

Tables 3 and 4 report descriptive statistics for the FG speech-day samples of the Federal Reserve and the Sveriges Riksbank, respectively. The four tone indices (Odyssean, Delphic, Negative, and Uncertainty) were reported in standardized form to facilitate comparability across dimensions. accordingly, each standardized tone index was centred around zero with a standard deviation of one within each bank sample.The Federal Reserve FG sample contained 194 observations, whereas the Riksbank sample contained 38 observations; consequently, banks comparisons were interpreted as indicative patterns rather than precise countries differences.

For the Fed (Table 3), the mean absolute daily change in the 2-year yield, $|\Delta 2Y|$, was 3.144 basis points, compared with 4.072 basis points for the 5-year yield, $|\Delta 5Y|$. Both distributions were strongly right-skewed (skewness of 3.062 and 2.167, respectively) and exhibited high kurtosis (15.007 and 9.811), indicating that most FG speech-days coincided with small or moderate yield moves while a small fraction of days featured very large reactions. This pattern was reflected in the upper-tail statistics: the 99th percentiles (p99) were 22.210 basis points for $|\Delta 2Y|$and 18.280 basis points for $|\Delta 5Y|$, with maxima of 26 and 24 basis points. At the lower tail, the minimum and the 1st percentile (p1) were 0.000 for both maturities, showing that some FG speech-days were associated with negligible absolute changes. Overall, these moments characterized an asymmetric and heavy-tailed distribution.

Regarding tone dispersion, the standardized indices displayed wide supports: Odyssean tone ranged from −1.803 to 2.638, Delphic tone from −2.278 to 3.165, Negative tone from −2.149 to 3.422, and Uncertainty tone from −2.042 to 4.127.The Head dummy averaged 0.469, implying that 46.9% of FG speech-days were delivered by the Chair, whereas the Covid dummy averaged 0.113, indicating that 11.3% of observations fell within the Covid period. Finally, $\log(\mathrm{GapDays})$ had a mean of 0.272 (standard deviation 0.488) and ranged from 0.000 to 1.386, documenting heterogeneity in the time elapsed between consecutive FG speech-days.

For the Riksbank (Table 4), the mean absolute daily changes on FG speech-days were 1.679 basis points for $|\Delta 2Y|$and 2.820 basis points for $|\Delta 5Y|$. The yield-change distributions were again right-skewed and heavy-tailed (skewness of 1.979 and 2.112; kurtosis of 7.156 and 8.293). The lower tail remained close to zero (p1 of 0.026 for $|\Delta 2Y|$and 0.137 for $|\Delta 5Y|$), while upper-tail values were sizable (p99 of 6.867 and 12.960, respectively). The standardized tone indices also exhibited wide supports despite the smaller sample. The Head dummy averaged 0.289 and the Covid dummy averaged 0.105, so approximately 29% of Riksbank FG speech-days were delivered by the Governor and about 11% occurred during the Covid period. Finally, $\log(\mathrm{GapDays})$ had a mean of 0.292 (standard deviation 0.506) and ranged from 0.000 to 1.609, indicating heterogeneous spacing of FG speech-days even within this smaller sample.

Taken together, Tables 3 and 4 shows that absolute yield reactions on FG speech-days were highly asymmetric and concentrated in right-tail moments, and that the tone indices varied sufficiently within the samples to support the regression hypotheses reported in the next sections.

Table 3: Descriptive Statistics — Federal Reserve (FG speech-day sample)

Variables	Obs	Mean	Std. Dev.	Min	Max	p1	p99	Skew.	Kurt.
$\lvert \Delta 2Y \rvert$	194	3.144	4.099	.000	26.000	.000	22.210	3.062	15.007
$\lvert \Delta 5Y \rvert$	194	4.072	3.810	.000	24.000	.000	18.280	2.167	9.811
Odyssean tone	194	.000	1.000	-1.803	2.638	-1.803	2.449	.441	2.808
Delphic tone	194	-.000	1.000	-2.278	3.165	-1.802	2.357	.294	2.753
Negative tone	194	.000	1.000	-2.149	3.422	-1.956	2.560	.711	3.696
Uncertainty tone	194	.000	1.000	-2.042	4.127	-1.914	2.686	.611	4.138
Head	194	.469	.500	.000	1.000	.000	1.000	.125	.995
Covid	194	.113	.318	.000	1.000	.000	1.000	2.458	7.081
log(GapDays)	194	.272	.488	.000	1.386	.000	1.386	1.278	2.729

Note: p1 and p99 denote the 1st and 99th percentiles. Kurt. reports Pearson kurtosis.

Table 4: Descriptive Statistics — Sveriges Riksbank (FG speech-day sample)

Variables	Obs	Mean	Std. Dev.	Min	Max	p1	p99	Skew.	Kurt.
$\lvert \Delta 2Y \rvert$	38	1.679	1.629	.000	7.200	.026	6.867	1.979	7.156
$\lvert \Delta 5Y \rvert$	38	2.820	2.974	.100	13.700	.137	12.960	2.112	8.293
Odyssean tone	38	-.000	1.000	-2.760	1.910	-2.429	1.907	-.183	3.658
Delphic tone	38	.000	1.000	-1.839	2.600	-1.803	2.338	.235	3.165
Negative tone	38	.000	1.000	-1.559	2.603	-1.554	2.538	.778	3.545
Uncertainty tone	38	.000	1.000	-2.250	2.688	-2.195	2.484	.169	4.240
Head	38	.289	.460	.000	1.000	.000	1.000	.967	1.873
Covid	38	.105	.311	.000	1.000	.000	1.000	2.679	8.464
log(GapDays)	38	.292	.506	.000	1.609	.000	1.420	1.286	2.946

Note: p1 and p99 denote the 1st and 99th percentiles. Kurt. reports Pearson kurtosis.

4.1.2 Tone evolution (all speeches)

To document how Odyssean and Delphic tone evolved over 2012–2024, the analysis first examined the full speech corpus (765 speeches for the Federal Reserve and 109 for the Riksbank) before comparing these patterns with the FG-only samples in the following subsection. Figures 1 and 2 displayed the annual evolution of Odyssean and Delphic tone for the Fed and the Riksbank, respectively, with both tone dimensions plotted in each figure. Tone was expressed as the number of tone words per 1,000 words, which normalized for variation in speech length and allowed level differences between Odyssean and Delphic tone to be interpreted directly.

For the Fed, Delphic tone remained consistently higher than Odyssean tone throughout 2012–2024. Delphic tone increased from the early years to a clear mid-sample peak around 2016, declined toward 2020, and then recovered in the later years, ending at 19.19 in 2024. Odyssean tone was comparatively stable, fluctuating in a narrower band; it increased modestly around 2020–2021, declined into 2023, and ended at 5.52 in 2024. Overall, the Federal Reserve all-speeches figure was characterized by a high Delphic baseline with visible medium-run swings alongside a lower and more stable Odyssean series.

For the Riksbank, Delphic tone likewise remained above Odyssean tone across the full period but followed a different time profile. Delphic tone started high in 2012, declined materially through the mid-to-late 2010s, reached a low point around 2021, and then rose sharply in 2022–2023 before easing slightly, ending at 16.63 in 2024. Odyssean tone was lower and more stable than the Delphic series, declining from 2012 to 2014, remaining relatively flat through the mid-2010s, increasing around 2020–2021, and ending at 5.45 in 2024. These figures were presented as descriptive context based on the full speech sets and were not restricted to the FG speech-day estimation samples.

Figure 1 : Fed tone over time (all speeches)

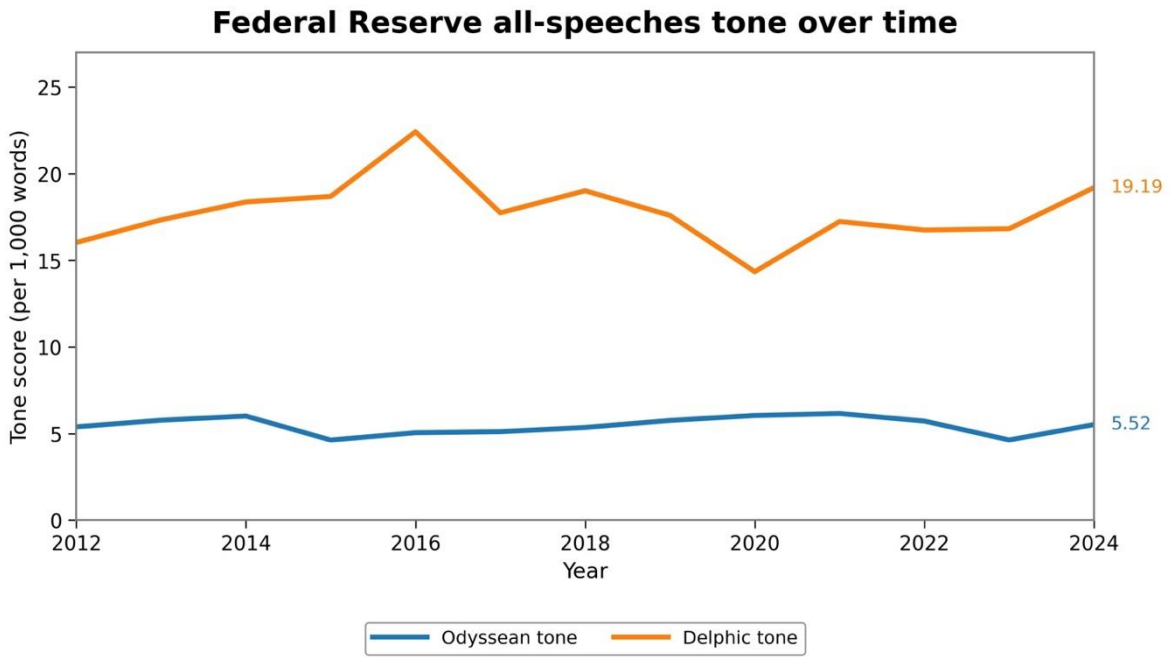
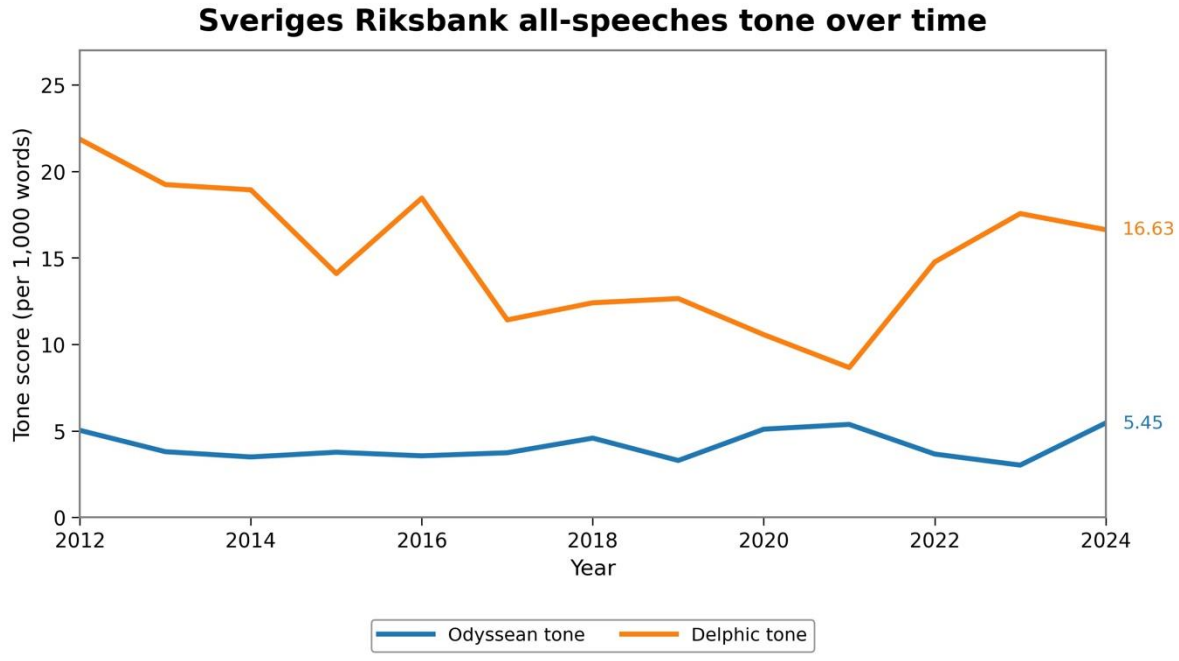


Figure 2 : Riksbank tone over time (all speeches)



4.1.3 Tone evolution (FG speech-days)

Compared with the all-speeches graphs, restricting the series to FG speeches produced sharper year-to-year movements, consistent with the fact that the annual means were computed from a smaller and more specific set of communication events.

For the Fed FG set (Figure 3), Delphic tone remained consistently above Odyssean tone over 2012–2024. Delphic tone started slightly above 20 words per 1,000 in 2012, rose sharply in the mid-2010s (above 25 around 2014 and 2016), and then declined gradually into 2020, when it reached roughly 18. It subsequently increased to a local high around 2021 (close to 24), remained elevated in 2022 (above 22), fell sharply in 2023 (to about 15), and partially recovered to 18.23 in 2024. Odyssean tone moved within a narrower range, mostly between 5 and 6 words per 1,000, with a gradual increase into the early 2020s and a local peak near 7.5 around 2021. It then declined through 2022–2023 and ended at 4.93 in 2024.

For the Riksbank FG set (Figure 4), Delphic tone again exceeded Odyssean tone throughout the period but displayed larger moves. Delphic tone was high in the early years (above 20 with a peak above 25 in 2014), then dropped sharply in 2015 (to around 10), recovered in 2016 (close to 19), and then showed a decreasing trajectory into the early 2020s, reaching a dip around 2021 (below 10). From 2022, it rose sharply, reaching 18.52 in 2024. Odyssean tone varied within a narrower range but increased over time. It was around 3–4 words per 1,000 in 2013–2014, rose gradually through the late 2010s, increased further around 2020–2021 (to roughly 6), decreased in 2022, and then rose again in 2023–2024, ending at 6.79.

Figure 3 : Fed tone over time (FG speeches only)

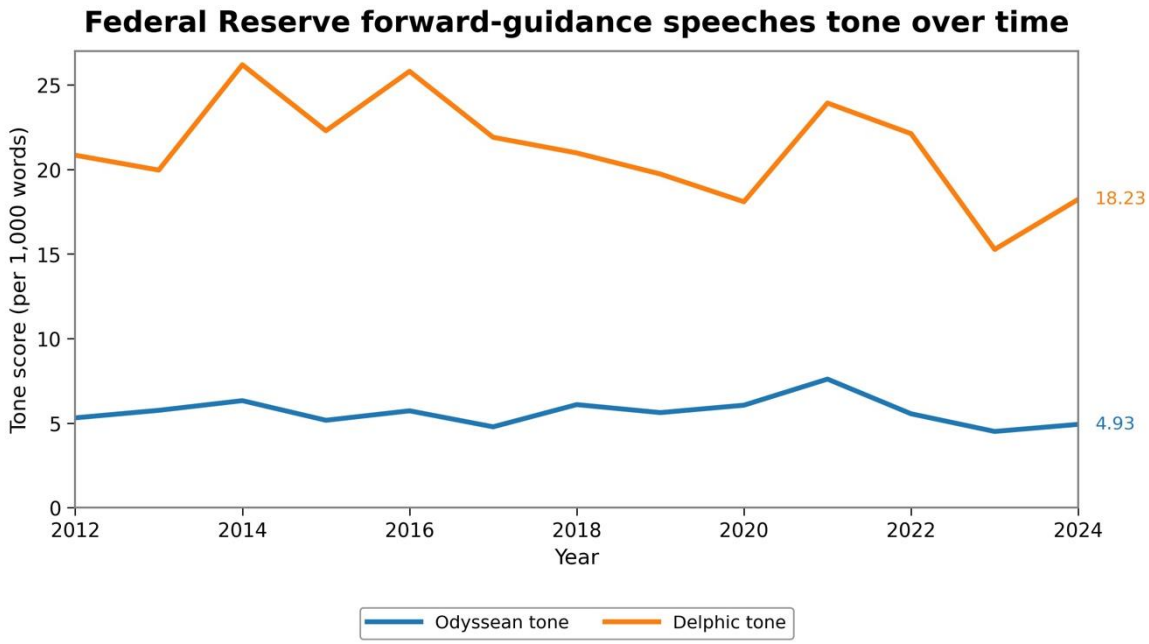
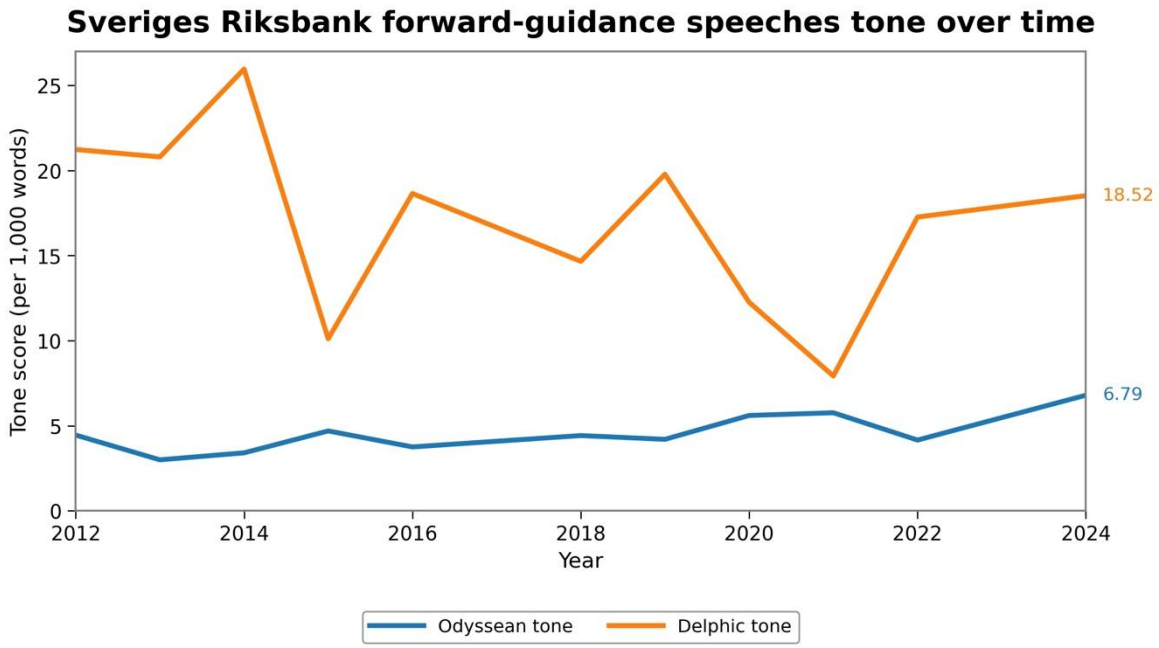


Figure 4 : Riksbank tone over time (FG speeches only)



Overall, the descriptive section documented clear time variation in FG tone measures and heavy-tailed distributions of yield reactions on FG speech-days. Section 4.2 therefore moved from description to inference by estimating the event-style regression models and evaluating the empirical support for the stated hypotheses.

4.2 Regression results

In this section, the empirical results for the association of FG communication tone and absolute yield changes are analyzed. The analysis first focuss on the Fed baseline models, where absolute daily 2- and 5-year yield changes on FG speech-days were regressed on standardized tone indices and a small set of controls, before turning to interaction and robustness checks and, finally, presenting the Riksbank case.

All regressions in this section were estimated by ordinary least squares with HC3 robust standard errors, in order to obtain inference that was robust to heteroskedasticity in daily yield changes. Variance inflation factors (VIFs) for all regressors remained well below the conventional critical threshold of 10 (Wooldridge, 2009), indicating that multicollinearity was far from levels usually considered problematic in linear regressions (see Appendix 7). Unless otherwise stated, statistical inference relied on two-tailed tests at standard significance levels, even when the hypotheses were directional ($\beta > 0/\ \beta < 0$); this conservative choice ensured a uniform decision rule across all coefficients and prevented cases from being over-interpreted.

4.2.1 Fed baseline results

Table 5 reports the baseline regression results for the Fed , where the dependent variable was the absolute daily change in government bond yields on FG speech-days. The four specifications used, respectively, same-day absolute 2-year yield changes ($|\Delta 2Y|$), next-day absolute 2-year changes, a trimmed version of $|\Delta 2Y|$ excluding the top 1% extreme observations, and same-day absolute 5-year yield changes ($|\Delta 5Y|$) as a simple maturity comparison.

The models had low explanatory power. R^2 ranged from 0.048 and 0.087 , while adjusted R^2 (more conservative) ranged from 0.012 to 0.053 across the four specifications. The HC3-robust F-test evaluated the joint null hypothesis that all slope coefficients were simultaneously equal to zero ($H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$). In Table 5, the corresponding model p-values ranged from 0.0001 to 0.021, so the joint null hypothesis was rejected in every specification. This indicated that, taken together, the regressors had statistically detectable explanatory power for the dependent variable.

Because all tone variables were standardized, each coefficient was interpreted as the change in expected absolute yield movement (in bp) associated with a one-standard-deviation change in the corresponding index, conditional on the other covariates. Odyssean tone coefficients were small and never statistically significant: -0.086 (0.267) for same-day $|\Delta 2Y|$, 0.020 (0.236) for next-day $|\Delta 2Y|$, -0.253 (0.202) in the trimmed 2-year regression, and -0.184 (0.268) for $|\Delta 5Y|$, all without significance stars at the conventional levels. Delphic tone coefficients followed a similar pattern, with estimates of -0.067 (0.428), -0.276 (0.329), 0.125 (0.371), and 0.149 (0.398) across Columns, again without statistical significance. In all cases, p-values for Odyssean and Delphic tone exceeded 0.10; therefore, the null hypotheses $H_0: \beta_1 = 0$ and $H_0: \beta_2 = 0$ could not be rejected in any baseline specification, and hypotheses H1 and H2 were not supported.

The sentiment control variables delivered more nuanced evidence. Negative tone was estimated with a negative sign in all columns and was statistically significant in the same-day and trimmed models: -0.640 (0.256) in Column (1) and -0.469 (0.214) in Column (3), both significant at the 5% level (two-tailed). In the next-day and 5-year regressions, Negative tone coefficients were -0.253 (0.192) and -0.263 (0.250), respectively, and these estimates were not statistically significant. Accordingly, $H_0: \beta_3 = 0$ was rejected in Columns (1) and (3) but not in Columns (2) and (4); however, the estimated sign was negative rather than the positive sign stated in H3, so the directional content of H3 was not supported.

For the Uncertainty tone index, coefficients were close to zero with large standard errors in all four specifications : 0.092 (0.464), -0.184 (0.261), -0.039 (0.434), and 0.028 (0.429) respectively, and all corresponding p-values exceeded 0.10. Therefore, the null hypothesis $H_0: \beta_4 = 0$ could not be rejected in any specification, and H4 was not supported.

The control variables indicated which aspects of the communication variables exhibited statistically detectable associations with $|\Delta Y|$. The Head speaker dummy was negative in each specification but was statistically significant at the 5% level only in the same-day 2-year regression, where the coefficient was -1.219 (0.611). In the other models, the Head estimates were -0.592 (0.454), -0.721 (0.502), and -0.901 (0.567), none reaching conventional significance levels. The Covid period dummy was uniformly negative and statistically significant in all four specifications, with coefficients of -2.904 (0.601) in Column (1), -1.341 (0.750) in Column (2), -2.375 (0.453) in Column (3), and -2.429 (0.638) in Column (4). Finally, $\log(\text{GapDays})$ was negative and statistically insignificant in all columns, so the null of no gap effect could not be rejected.

Taken together, the baseline Fed regressions showed that Odyssean and Delphic tone did not display statistically detectable associations with the magnitude of yield reactions to FG speeches. The data did not provide sufficient evidence to reject the null hypotheses associated with H1, H2, and H4, and while a statistically significant relationship involving Negative tone was detected in the 2-year same-day and trimmed specifications, the estimated sign was opposite to the directional prediction in H3.

Table 5: Regression results — Federal Reserve (baseline specifications)

Dependent variable: absolute daily yield change on FG speech-days				
	(1) $\|\Delta 2Y\|$	(2) Next-day $\|\Delta 2Y\|$	(3) Trimmed $\|\Delta 2Y\|$	(4) $\|\Delta 5Y\|$
Odyssean tone (std.)	-0.086	0.020	-0.253	-0.184
	(0.267)	(0.236)	(0.202)	(0.268)
Delphic tone (std.)	-0.067	-0.276	0.125	0.149
	(0.428)	(0.329)	(0.371)	(0.398)
Negative tone (std.)	-0.640**	-0.253	-0.469**	-0.263
	(0.256)	(0.192)	(0.214)	(0.250)
Uncertainty (std.)	0.092	-0.184	-0.039	0.028
	(0.464)	(0.261)	(0.434)	(0.429)
Head speaker	-1.219**	-0.592	-0.721	-0.901
	(0.611)	(0.454)	(0.502)	(0.567)
COVID period	-2.904***	-1.341*	-2.375***	-2.429***
	(0.601)	(0.750)	(0.453)	(0.638)
log(GapDays)	-0.666	-0.060	-0.391	-0.489
	(0.526)	(0.546)	(0.485)	(0.552)
Constant	4.227***	3.085***	3.635***	4.904***
	(0.622)	(0.357)	(0.472)	(0.533)
Observations	194	194	192	194
R^2	0.087	0.048	0.078	0.057
Adj. R^2	0.053	0.012	0.043	0.021
Wald F (HC3)	5.019	2.436	5.759	2.524
Prob $> F$	0.000	0.021	0.000	0.017

Notes: HC3 robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

4.2.2 Fed interaction effects

This section examines whether the market sensitivity to forward-guidance tone varied across speaker authority (Head) and crisis conditions (Covid). Table 6 reports regressions in which standardized Odyssean and Delphic indices were interacted with a Head speaker dummy (Column 1) and a Covid-period dummy (Column 2). The interaction coefficients in Table 6 tested whether the tone–yield relationship differed across regimes (H5-H7), while Table 7 reports the implied marginal effects , which

is the actual tone–yield association within each regime, computed as linear combinations of the main and interaction coefficients.

Column (1) of Table 6 tested H5 and H6, which supposed that the association between tone and absolute yield changes differed when the speaker was the Fed Chair. The main tone coefficients captured the marginal association for non-Head speeches (Head = 0): Odyssean tone was 0.105 (0.490) and Delphic tone was 0.342 (0.573), neither statistically significant. The interaction terms Odyssean tone × Head and Delphic tone × Head were -0.338 (0.600) and -0.777 (0.535), respectively, and were also statistically insignificant.

In Table 7, Column (1) reports the implied marginal effects. For Odyssean tone, the effect for non-Chair speeches was 0.105 (0.490) and for Chair speeches was -0.233 (0.301). For Delphic tone, the respective effects were 0.342 (0.573) and -0.435 (0.445). None of these regime-specific estimates was statistically significant. The null hypotheses $H_0: \theta_1 = 0 \; and \; H_0: \theta_2 = 0$ could not be rejected, and H5 and H6 were not supported.

The Head speaker main effect was -1.262 (0.622), significant at the 5% level. The control variables were consistent with the baseline specifications: Negative tone was -0.676 (0.258), significant at the 1% level; Uncertainty was 0.052 (0.469), not significant; Covid period was -3.180 (0.771), significant at the 1% level; and log(GapDays) was -0.723 (0.541), not significant. Model R^2 was 0.100, adjusted R^2 was 0.056, and the model p-value was < 0.001.

Column (2) of Table 6 tested H7, which supposed whether the tone–yield association differed between pre- Covid and Covid regimes. The main tone coefficients represents the pre-pandemic baseline: Odyssean tone was 0.013 (0.330) and Delphic tone was -0.063 (0.451), neither statistically significant. The interaction terms were -0.574 (0.477) for Odyssean tone × Covid and 0.038 (0.530) for Delphic tone × Covid, and both were statistically insignificant.

Table 7, Column (2) reports the implied marginal effects. For Odyssean tone, the pre- Covid effect was 0.013 (0.330) and the during- Covid effect was -0.561 (0.282), the latter statistically significant at the 5% level. For Delphic tone, the pre- Covid effect was -0.063 (0.451) and the during- Covid effect was -0.025
(0.509), neither significant. The Covid -period marginal effect for Odyssean tone was significant, but the corresponding interaction coefficient in Table 6 was not.

The null hypotheses $H_0: \varphi_1 = 0 \; and \; H_0: \varphi_2 = 0$ could not be rejected, and H7 was not supported. The control variables followed similar patterns: Negative tone was -0.651 (0.259), significant at the 5% level; Uncertainty was 0.072 (0.466), not significant; Head speaker was -1.310 (0.657), significant at the 5% level; Covid period was -2.790 (0.617), significant at the 1% level; and log(GapDays) was -0.620 (0.531), not significant. Model R^2 was 0.089, adjusted R^2 was 0.045, and the model p-value was < 0.001, meaning the regressors taken together have statistically significant explanatory power for the dependent variable.
The interaction terms for both Head and Covid were statistically insignificant, and H5, H6, and H7 were not supported.

Table 6: Regression results — Federal Reserve (heterogeneity)

Dependent variable: absolute daily yield change on FG speech-days		
	(1) Head interaction $\lvert\Delta 2Y\rvert$	(2) COVID interaction $\lvert\Delta 2Y\rvert$
Odyssean tone (std.)	0.105	0.013
	(0.490)	(0.330)
Delphic tone (std.)	0.342	-0.063
	(0.573)	(0.451)
Odyssean tone (std.) × Head	-0.338	
	(0.600)	
Delphic tone (std.) × Head	-0.777	
	(0.535)	
Odyssean tone (std.) × COVID		-0.574
		(0.477)
Delphic tone (std.) × COVID		0.038
		(0.530)
Negative tone (std.)	-0.676***	-0.651**
	(0.258)	(0.259)
Uncertainty (std.)	0.052	0.072
	(0.469)	(0.466)
Head speaker	-1.262**	-1.310**
	(0.622)	(0.657)
COVID period	-3.180***	-2.790***
	(0.771)	(0.617)
log(GapDays)	-0.723	-0.620
	(0.541)	(0.531)
Constant	4.333***	4.264***
	(0.682)	(0.635)
Observations	194	194
R^2	0.100	0.089
Adj. R^2	0.056	0.045
Robust F-statistic (HC3)	4.072	5.749
P-value model	0.000	0.000

Notes: HC3 robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

Table 7: Implied marginal effects from interaction models — Federal Reserve

	(1) Head interaction	(2) Covid interaction
Odyssean tone (std.)		
Effect in reference regime (Head=0 / Pre-Covid)	0.105	0.013
	(0.490)	(0.330)
Effect in alternative regime (Head=1 / During Covid)	-0.233	-0.561**
	(0.301)	(0.282)
Delphic tone (std.)		
Effect in reference regime (Head=0 / Pre-Covid)	0.342	-0.063
	(0.573)	(0.451)
Effect in alternative regime (Head=1 / During Covid)	-0.435	-0.025
	(0.445)	(0.509)
Sample size matches the interaction regressions		

Notes: Entries report implied marginal effects of tone computed as linear combinations of interaction-model coefficients. HC3 robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

4.2.3 Fed robustness checks

Table 8 reports robustness checks that assessed the sensitivity of the baseline results to alternative variable specifications. All four columns used the same dependent variable, the absolute daily 2-year yield change on FG speech-days, and used the HC3 robust standard errors. Column (1) reproduced the baseline specification from Table 5 for direct comparison. Columns (2), (3), and (4) altered the set of tone regressors to evaluate whether the baseline findings were sensitive to the particular choice of tone indices. All columns used 194 observations.

Column (2) excluded the Uncertainty index while retaining all other baseline regressors. The coefficients on Odyssean tone and Delphic tone were not statistically significant: -0.096 (0.259) and -0.004 (0.267), respectively. Negative tone remained statistically significant at the 1% level with a coefficient of -0.617 (0.234). The control variables followed the same pattern as in the baseline: Head speaker was -1.240 (0.635), significant at the 10% level; Covid period was -2.921 (0.619), significant at the 1% level; and log(GapDays) was -0.661 (0.513), not significant.

Column (3) excluded the Delphic tone index. Odyssean tone remained not statistically significant at -0.104 (0.252). Negative tone was -0.633 (0.249), significant at the 5% level, and Uncertainty was 0.045 (0.297), not significant. The Head speaker coefficient was -1.225 (0.613), significant at the 5% level, and the Covid period coefficient was -2.901 (0.595), significant at the 1% level.

Column (4) replaced the separate Odyssean and Delphic indices with a single Net forward guidance index (NetFG), computed from standardized tone indices as the difference between Odyssean and Delphic dimensions. The NetFG coefficient was -0.049 (0.292), not statistically significant. Negative tone remained negative and significant at the 5% level with a coefficient of -0.624 (0.260), while Uncertainty was 0.004 (0.351), not significant. The control variables followed the baseline results: Head speaker was -1.258 (0.602), significant at the 5% level; Covid period was -2.932 (0.601), significant at the 1% level; and log(GapDays) was -0.639 (0.505), not significant.

Finally , Model R^2 ranged from 0.086 to 0.087, and adjusted R^2 ranged from 0.053 to 0.058. The HC3 robust F-statistics yielded model p-values < 0.001 in all columns, confirming joint statistical significance of the regressors. Coefficient estimates and statistical inference remained stable across alternative tone specifications.

Table 8: Regression results — Federal Reserve (robustness)

Dependent variable: $\|\Delta 2Y\|$ (absolute daily yield change on FG speech-days)				
	(1) Baseline	(2) No Unc.	(3) No Delphic	(4) NetFG
Odyssean tone (std.)	-0.086	-0.096	-0.104	
	(0.267)	(0.259)	(0.252)	
Delphic tone (std.)	-0.067	-0.004		
	(0.428)	(0.267)		
Net forward guidance (std.)				-0.049
				(0.292)
Negative tone (std.)	-0.640**	-0.617***	-0.633**	-0.624**
	(0.256)	(0.234)	(0.249)	(0.260)
Uncertainty (std.)	0.092		0.045	0.004
	(0.464)		(0.297)	(0.351)
Head speaker	-1.219**	-1.240*	-1.225**	-1.258**
	(0.611)	(0.635)	(0.613)	(0.602)
COVID period	-2.904***	-2.921***	-2.901***	-2.932***
	(0.601)	(0.619)	(0.595)	(0.601)
log(GapDays)	-0.666	-0.661	-0.666	-0.639
	(0.526)	(0.513)	(0.520)	(0.505)
Constant	4.227***	4.237***	4.229***	4.241***
	(0.622)	(0.638)	(0.623)	(0.620)
Observations	194	194	194	194
R^2	0.087	0.087	0.087	0.086
Adj. R^2	0.053	0.058	0.058	0.057
Robust F-statistic (HC3)	5.019	5.878	5.947	5.622
P-value model	0.000	0.000	0.000	0.000

Notes: HC3 robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

4.2.4 Riksbank baseline results (comparative case)

The Riksbank analysis was treated as a comparative benchmark rather than as a fully symmetric counterpart to the Fed analysis due to its small sample size (38 observations). Accordingly, only the baseline specification was estimated, and no interaction or robustness extensions were conducted.

Table 9 reports the Riksbank baseline regression using the same specification structure as Section 4.2.1. Model analysis revealed a strong divergence from the Fed results. R^2 was equal to 0.355 for the 2-year yield and 0.248 for the 5-year yield, which was considerably higher than the Fed baseline R^2 of approximately 0.09. Adjusted R^2 was equal to 0.204 and 0.072, respectively. However, the HC3 robust F-tests did not reject the joint null hypothesis that all slope coefficients were equal to zero, with model p-values of 0.191 for |Δ2Y| and 0.167 for |Δ5Y|. This contrasts sharply with the Fed sections, where every baseline specification yielded significant F-tests ($p < 0.001$) . The combination of high R^2 but non-significant F-tests may reflect limited degrees of freedom (38 observations, 7 slope coefficients) and potentially unstable coefficient estimates in the small sample.

For the 2-year yield (Column 1), Odyssean tone was positive and statistically significant at the 10% level: 0.512 (0.311), indicating that stronger commitment-based language was associated with larger absolute yield movements. This positive association goes against the theoretical predictions that Odyssean guidance should reduce market volatility through credible commitment. Delphic tone was 0.723 (0.475), not statistically significant. Particularly, the Odyssean coefficient more than doubled at the 5-year maturity: 1.185 (0.605), significant at the 5% level, suggesting that the tone-volatility relationship was stronger at longer maturities. Delphic tone remained insignificant: 1.327 (0.963).

Regarding the sentiment controls, Negative tone was positive and statistically significant at conventional levels for both maturities: 0.944 (0.406) for the 2-year yield and 1.179 (0.695) for the 5-year yield, unlike the Fed results, where Negative tone was consistently negative (Table 5). Uncertainty was negative and statistically significant at the 5% level in both specifications: -1.046 (0.464) for the 2-year yield and -1.768 (0.898) for the 5-year yield.

The Head speaker dummy was not statistically significant in either specification: 0.108 (0.740) for the 2-year yield and 0.258 (1.402) for the 5-year yield. In contrast to the Fed, where the Head speaker coefficient was consistently negative and significant, the Riksbank showed no speaker-authority effect. The Covid period dummy was negative and statistically significant at the 10% level for the 2-year yield: -1.626 (0.918). The log(GapDays) coefficient was not significant for the 2-year yield but was negative and significant at the 5% level for the 5-year yield: -1.409 (0.618), which means that when the time since the previous FG-speech is larger, the absolute 5-year yield changes tend to be smaller.

Overall, the Riksbank coefficient patterns diverged from the Fed baseline across several dimensions.However, the small sample size and the non-rejection of the model-level F-tests constrained statistical inference. Given these constraints, the results were reported as exploratory comparative evidence that suggests potential heterogeneity in tone transmission mechanisms across central banks, even though formal hypothesis testing was not conducted.

Table 9: Regression results — Riksbank (baseline)

Dependent variable: absolute daily yield change on FG speech-days		
	(1) $\lvert\Delta 2Y\rvert$	(2) $\lvert\Delta 5Y\rvert$
Odyssean tone (std.)	0.512*	1.185**
	(0.311)	(0.605)
Delphic tone (std.)	0.723	1.327
	(0.475)	(0.963)
Negative tone (std.)	0.944**	1.179*
	(0.406)	(0.695)
Uncertainty (std.)	-1.046**	-1.768**
	(0.464)	(0.898)
Head speaker	0.108	0.258
	(0.740)	(1.402)
COVID period	-1.626*	-1.763
	(0.918)	(1.859)
log(GapDays)	-0.365	-1.409**
	(0.357)	(0.618)
Constant	1.925***	3.342***
	(0.362)	(0.667)
Observations	38	38
R^2	0.355	0.248
Adj. R^2	0.204	0.072
Robust F-statistic (HC3)	1.545	1.623
P-value model	0.191	0.167

Notes: HC3 robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

5. Discussion

5.1 Summary of Key Findings

The empirical results highlight three principal elements about speech-based forward guidance in the Fed and the Riksbank over 2012–2024.

Firstly, for the Fed, results do not support a systematic relationship between speech-level Odyssean or Delphic tone and the magnitude of same-day yield movements on forward-guidance speech days. Across the baseline specification and all robustness variants, the estimated coefficients on the standardized Odyssean index are economically small and statistically not different from zero . The same conclusion holds for the standardized Delphic index and it persists under alternative specifications and a trimmed sample. Interpreted conservatively, within the empirical model adopted here, variation in odyssean/commitment or in Delphic/forecast vocabulary in Fed speeches does not significantly lead to larger or smaller absolute daily changes in 2-year or 5-year government bond yields.

Secondly, in the Fed sample, the most stable associations come from control variables rather than from the Odyssean or Delphic indices, and these associations exhibit a clear maturity gradient. Negative tone is consistently related to smaller absolute yield changes at the 2-year maturity, with a statistically significant negative coefficient in both the baseline and trimmed models (e.g., −0.640, $p < 0.05$ in the baseline; Table 5). However, this relationship weakens substantially at the 5-year maturity (−0.263, $p > 0.10$), where the coefficient loses statistical significance. Similarly, Chair speeches are associated with smaller absolute yield changes at 2-year maturities (−1.219, $p < 0.05$; Table 5) but this effect becomes marginal and statistically insignificant at 5-year yields (−0.901, $p > 0.10$). This maturity gradient is consistent with forward-guidance theory: shorter-maturity yields, which more directly reflect expectations about the near-term policy rate path, respond more sharply to communication features such as speaker authority and sentiment tone (Gürkaynak et al., 2005; Swanson, 2017). By contrast, the Covid-19 period dummy is the only variable that exhibits robust significance across both maturities (−2.904, $p < 0.01$ at 2-year; −2.429, $p < 0.01$ at 5-year; Table 5), indicating that the pandemic regime compressed yield volatility uniformly across the curve, likely reflecting the dominance of concurrent policy interventions and extreme macro shocks over incremental speech-level communication. The Chair × tone interaction terms are not statistically significant, implying that the Chair effect does not vary systematically with Odyssean or Delphic tone. Similarly, although the Odyssean × Covid interaction coefficient is not statistically significant, the implied marginal effect of Odyssean tone during Covid is negative and significant (−0.561, $p < 0.05$; Table 7), indicating that higher Odyssean tone is associated with smaller absolute yield changes in the pandemic regime. This inference is conditional on the interaction specification, as it is based on a linear combination of estimated coefficients rather than on the interaction coefficient alone.

Finally, the Riksbank results are suggestive but do not allow definitive conclusions due to severe sample-size constraints (n = 38). Individual Odyssean coefficients are positive and reach conventional significance thresholds in both maturity specifications: 0.512 ($p < 0.10$) for 2-year yields and 1.185 ($p < 0.05$) for 5-year yields (Table 9). Strikingly, the Odyssean coefficient more than doubles in magnitude when moving from 2-year to 5-year maturities, and the level of statistical significance strengthens from marginal to conventional. This maturity gradient is the opposite of the pattern observed in the Fed sample and contradicts the theoretical prediction that forward guidance should primarily affect shorter-maturity yields. One interpretation is institutional: within the Riksbank's established Delphic framework, which is built on conditional rate-path forecasts and scenario-based transparency (Moessner et al., 2014; Breman, 2025), the use of commitment-oriented Odyssean language is rare and therefore potentially interpreted by markets as a signal of fundamental regime change rather than

incremental guidance. If markets view such departures as newsworthy precisely because they deviate from the bank's established forecast-based norm, the effect may be amplified at medium-term maturities, where forward guidance is expected to exert less influence under standard theory. However, the model-level F-tests do not reject joint non-significance (Table 9), and the small sample limits statistical power and increases sensitivity to influential observations. Consequently, the Riksbank evidence is best interpreted as exploratory. It raises the possibility that the relationship between commitment-like language and volatility differs systematically across institutional contexts and across maturities, but it does not establish a robust contrast. Further research with larger samples or alternative identification strategies would be required to confirm whether the observed maturity amplification reflects a genuine institutional transmission mechanism or a statistical artifact of the small-sample setting.

5.2 Interpretation of Results

5.2.1 Odyssean and Delphic tone for the Fed

The absence of statistically significant associations between Odyssean/Delphic tone and absolute daily yield changes in the Federal Reserve sample should be interpreted against two features of the empirical design.

First, the dependent variable is the magnitude of yield changes ($|\Delta Y|$), which captures volatility rather than the direction of policy news. Second, as discussed in the literature review, theoretical predictions for magnitude are not unambiguous: credible commitment can reduce policy uncertainty and dampen market movements (Campbell et al., 2012) but the same communication may also trigger immediate repricing by concentrating information or resolving uncertainty (Nakamura & Steinsson, 2013; Goy et al., 2018). The empirical "null" therefore does not contradict a single sharp theoretical sign restriction. Rather, it indicates that within daily data and speech-level lexicon measures, any average association is either weak, or too noisy to detect.

A first interpretation is that the informational content of speeches may be dominated by the broader communication environment, reducing the marginal contribution of speech-level wording. Recent critiques of the Fed's post-2020 framework emphasize that market beliefs did not adjust smoothly to guidance language when incoming data and policy actions shifted the feasible policy path. Romer and Romer (2024) argue that odyssean guidance did not anchor expectations as intended, while Bauer et al. (2024) document limited updating by markets and professional forecasters in late 2021 and a more discrete repricing once policy action clearly diverged from prior guidance. If market participants place greater weight on realized policy actions, formal FOMC communications, and the evolving macro data flow than on the lexical tone of individual speeches, then variation in speech tone may not translate into detectable differences in daily yield volatility. This reading is consistent with Moessner et al. (2015), who emphasize a persistent tension between the theoretical appeal of binding commitments and policymakers' preference for discretion, which can compromise the credibility channel that would otherwise amplify Odyssean language.

A second interpretation is that daily yield movements are an imprecise proxy for the communication channel that FG is intended to operate through. The literature on publishing interest-rate paths suggests that communication can influence expectations beyond the immediate announcement window and that markets treat published paths as conditional forecasts rather than binding commitments (Moessner et al., 2014). If the underlying channel operates through a gradual revision of expectations rather than immediate same-day "fluctuations", an empirical design based on $|\Delta Y|$ at the daily frequency may understate effects. In other words, the null finding may stem from a timing mismatch: forward guidance can influence expectations gradually, whereas the dependent variable is designed to capture only same-day absolute yield movements.

A third interpretation is that the null result is consistent with "mixed signals" , as emphasized by the information-effect literature. Central bank communication may simultaneously transmit a policy

signal about future actions and an information signal about the underlying state of the economy (Nakamura & Steinsson, 2013; Andrade & Ferroni, 2021). Under this view, stronger Odyssean language can be interpreted both as greater intended accommodation and as evidence that the central bank perceives conditions as weaker or riskier than markets expected. These components can push yields in opposing ways and may therefore offset each other in absolute terms, reducing the observable relationship between lexical tone and same-day yield volatility. Because the empirical design does not separately identify policy and information components, the null finding should be interpreted as compatible with such offsetting effects rather than as evidence that communication is irrelevant

5.2.2 Negative tone

Negative tone is the only tone dimension that is consistently statistically significant in the Federal Reserve sample, but its estimated sign contradicts the usual expectation that negative language should heighten anxiety and amplify market volatility (Loughran & McDonald, 2012). The appropriate interpretation is therefore not that negative tone mechanically "reduces uncertainty", but rather that within this dataset and empirical specification, speech days with higher measured negative vocabulary are associated with smaller absolute same-day yield changes.

A modest explanation follows directly from the construction of the dependent variable. Because the outcome is the absolute yield change, any reaction that is directionally systematic and relatively predictable can register as a lower value of $|\Delta Y|$. If negative language is primarily interpreted as reinforcing downside risks and, by implication, a more accommodative reaction function, yields may adjust in a relatively smooth and one-sided manner rather than through abrupt repricing. In such a setting, the market response can be economically significant in sign and direction, yet still translate into smaller measured "volatility" when volatility is proxied by $|\Delta Y|$.

Second, negative language may be used in contexts where policymakers attempt to manage expectations cautiously, and where communication is deliberately calibrated to avoid and limit abrupt repricing. In that case, the estimated coefficient likely captures the context and timing of such communications rather than a direct, causal "stabilizing" effect of negative wording itself. More Importantly, the present design cannot decide between these possibilities. Distinguishing them would require, at minimum, complementary analysis using signed yield changes and/or narrower event windows that isolate the speech from same-day confounding factors. Thus, the negative coefficient is best reported as a robust empirical association that challenges the simple "negative tone increases volatility" prediction, while remaining open about the limits of causal interpretation.

5.2.4 Head and Covid effects

The Chair and Covid coefficients indicate that authority and regime context is empirically consequential for yield volatility on forward-guidance speech days, even when tone indices are held constant. In the Fed sample, Chair speeches are associated with smaller absolute yield changes than speeches by other FOMC participants (Table 5). This is not technically inconsistent with the argument that Chair communication is influential (Swanson & Jayawickrema, 2024). It may instead indicate that Chair communication is more anticipated, more tightly coordinated with the broader policy signal, or more likely to occur in environments where market expectations are already aligned, thereby reducing surprise and daily repricing. A related interpretation is that, subject to the subset of days identified as forward-looking speech days, the chair's speeches may contain less new information for markets than other communication channels. In this case, the negative coefficient should be interpreted as reflecting greater predictability and closer coordination of messages, hence smaller absolute yield adjustments, rather than implying that the Chairman's speeches are generally “less important” in shaping expectations.
The Covid-period coefficient is large and negative (Table 5), indicating smaller absolute yield movements on forward-guidance speech days during the pandemic regime. This pattern is consistent with a period in which yields were heavily influenced by concurrent interventions and dominant macro shocks, potentially compressing the incremental role of speech-level tone in day-to-day repricing. The interaction results add a narrow nuance: the implied marginal effect of Odyssean tone during Covid is negative and statistically significant (Table 7), which is consistent with the view that commitment-like language may have been interpreted as reassurance during an extreme-uncertainty regime. However, because the interaction term itself is not statistically significant, the safest conclusion is that the evidence points to possible regime dependence rather than providing definitive proof of a distinct “Covid-only” Odyssean effect.

5.2.5 Riksbank results

The Riksbank regressions suggest a positive association between Odyssean tone and absolute yield changes in some specifications, but the evidentiary strength is limited by sample size and by the absence of joint model significance (Table 9). Accordingly, the appropriate interpretation is not that Odyssean language “increases volatility” in Sweden as a stable fact, but that the point estimates are consistent with a setting in which odyssean language is perceived as relatively significant (or potentially debatable) within a communication framework that is widely understood as conditional and forecast-based.

The literature on the Riksbank’s rate-path framework emphasizes that markets interpret published paths as conditional forecasts and continue to respond to macroeconomic data rather than treating the path as a binding commitment (Moessner et al., 2014), while early evidence suggests the publication of rate forecasts can influence medium-term yields (Andersson & Hofmann, 2009). In such a framework, explicit commitment-like vocabulary may be rarer and therefore more “newsworthy,” potentially coinciding with larger repricing episodes. Svensson’s (2015) discussion of the “leaning against the wind” episode further highlights that historical credibility and policy reversals can shape how markets interpret guidance-like signals. These mechanisms are consistent with the pattern of positive coefficients; however, given the statistical fragility of the Riksbank sample, the safest conclusion is that the results motivate institutional-heterogeneity hypotheses rather than establishing them.

5.3 Implications

5.3.1 Implications for forward-guidance transmission and measurement

A central implication of the findings is that speech-level indices with words counting may capture meaningful variation in communication style without necessarily mapping into same-day yield volatility at the daily frequency. In the Fed sample, the absence of robust tone effects alongside significant context variables suggests that market repricing on forward-guidance speech days is shaped less by lexical tone in isolation than by institutional features (who speaks) and macro-financial regimes (when they speak). This is consistent with the view that the effectiveness of guidance depends on coordination across communication channels and on the credibility of the broader framework, not merely on the linguistic intensity of commitment or forecast vocabulary (Moessner et al., 2015).
At the same time, the results do not imply that forward guidance is irrelevant. Rather, they indicate that within a design focusing on daily $|\Delta Y|$ and speech-level lexicon measures, the identifiable marginal association of Odyssean/Delphic tone is weak. This distinction matters for interpretation as the thesis informs the measurement and empirical identification of speech-based FG effects, not the full set of channels through which guidance can influence expectations and the term structure over longer horizons.
The maturity-specific patterns documented in Section 5.2.4 add a further layer of nuance to the interpretation of speech-based forward guidance. The finding that tone effects vary systematically across the 2-year and 5-year yield curve segments suggests that forward guidance may not operate uniformly across all maturities. This heterogeneity is particularly relevant for reconciling the "forward guidance puzzle" (Del Negro et al., 2012): if markets respond to communication tone differently at short versus medium-term horizons, empirical designs that focus exclusively on a single maturity may understate or mischaracterize the full transmission effect. Moreover, the institutional divergence between the Fed and the Riksbank reinforces the broader point that the effectiveness and transmission of guidance depend not only on lexical tone but also on the institutional framework and the established communication norms of each central bank.

5.3.2 Implications for the Odyssean versus Delphic debate

The descriptive dominance of Delphic over Odyssean tone, combined with episodic rather than monotonic increases in Odyssean language, is consistent with the institutional reluctance to adopt binding commitment technologies emphasized in the post-crisis literature (Moessner et al., 2015). While optimal-policy analyses at the ELB highlight the theoretical power of commitment (Eggertsson and Woodford, 2003; Woodford, 2012), the present evidence suggests that, at least in speeches, central banks did not move toward a persistently more Odyssean lexical profile over 2012–2024, and any episodic shifts were not reliably associated with larger or smaller same-day yield volatility.

This pattern aligns with recent policy debates that criticize rigid outcome-based guidance and emphasize the value of communicating reaction functions and uncertainty in a disciplined way (English & Sack, 2024; Romer & Romer, 2024; Bernanke, 2025). However, the thesis does not establish that Delphic approaches are superior; it shows that Delphic language remains more prevalent in the speech corpus and that speech-level tone indices do not robustly explain daily yield volatility in the Fed sample. Accordingly, this thesis contributes to the literature mainly through descriptive evidence and methodological insight, rather than by offering a controlling assessment of which guidance type is superior.

5.3.3 Implications for communication design and coordination

For communication practice, the key implication is that coherence, timing, and the identity of the communicator may matter at least as much as lexical tone for market volatility on forward-guidance speech days. In the Fed sample, the Chair and Covid indicators are systematically related to $|\Delta Y|$, whereas the Odyssean and Delphic indices are not. This is consistent with concerns that fragmented communication across speeches, projections, and formal statements can complicate inference about the policy signal and reduce the marginal information of any single speech (Bernanke, 2025).

For the Riksbank, the tentative evidence that Odyssean language coincides with larger absolute yield moves suggests that deviations from an established, conditional, forecast-based communication style may be interpreted as primary and potentially destabilizing. This implication is necessarily cautious given statistical power, but it reinforces a general principle: communication strategies are institution-specific, and the same linguistic signal may be received differently depending on the central bank's framework and credibility history (Svensson, 2015; Moessner et al., 2014).

5.4 Limitations

The analysis relies on 194 forward-guidance speeches for the Federal Reserve and 38 for the Riksbank. This imbalance is itself informative about communication intensity, but it limits comparability. With n = 38, the Riksbank regressions have low power to detect moderate associations, and the failure of the model F-tests to reject joint insignificance indicates that inference based on individual coefficients is fragile . The appropriate conclusion is therefore that the Riksbank results are exploratory.

A second sample-related limitation is the identification of "forward-guidance speeches" using a keyword-based filter. While this approach enhances conceptual focus like Benchimol et al. (2025) did, it may exclude speeches that convey guidance implicitly through discussion of the reaction function, risks, or the policy framework without using explicit FG-related keywords. Any keyword filter involves a compromise between precision and coverage. Accordingly, the results should be interpreted as belonging to an empirically identified subset of speeches rather than to the full universe of monetary policy communication.

Dictionary-based sentiment analysis is transparent and replicable, but it is limited in construct validity for the Odyssean/Delphic distinction. The approach is insensitive to syntax and pragmatic context (Mao et al., 2024) : the presence of commitment vocabulary does not guarantee that the sentence represents an unconditional policy commitment, and a limited negation heuristic cannot reliably detect conditional or hypothetical framing. Moreover, commitment and forecast language can refer to multiple policy domains (e.g., supervisory commitments versus rate-path guidance). As a result, the Odyssean and Delphic indices may capture broad rhetorical style rather than a clean measure of the theoretical constructs defined by Campbell et al. (2012).
In addition, central-bank language evolves, and institution-specific phrases can carry policy meaning that static dictionaries do not capture. The deliberate exclusion of ambiguous high-frequency modals (e.g., "will") reduces mechanical correlations with speech length but may also omit genuine commitment content in some contexts. Collectively, these features imply measurement error in the tone indices, which can attenuate estimated coefficients and reduce the probability of detecting true associations, particularly in a setting where daily yield changes are inherently noisy.

The regression design uses absolute daily yield changes, which are influenced by all information processed within the day, not solely by the speech. This "contamination" increases residual variance and reduces precision. Timing mismatches, such as speeches delivered after market close that are mapped onto the next trading day, introduce additional noise. Endogeneity is also a concern: speech timing and tone may respond to recent market conditions or policy uncertainty, and such reverse causality cannot be fully eliminated with the available controls. Consequently, the estimates should be interpreted as conditional associations rather than clean causal effects of tone.

The focus on 2-year yields is theoretically motivated, as forward guidance is expected to operate primarily through short-to-medium-term expectations. Nevertheless, even 2-year yields embed term-premium variation and broader risk conditions, and the analysis does not decompose yields into expectations versus term-premium components, limiting a tight mapping from coefficients to specific transmission channels. In the baseline results, 5-year yields are included as a comparative maturity check, which provides partial reassurance that the main conclusions are not driven by the exclusive choice of horizon. However, the full set of robustness analyses was implemented for the 2-year maturity only, as a deliberate choice to prioritize depth of specification checks at the horizon most directly linked to forward guidance. This choice improves internal consistency for the primary outcome but necessarily limits full comparability across maturities.

Finally, the sample spans heterogeneous regimes: post-crisis recovery, negative rates in Sweden, the Covid shock, and rapid normalization after 2022. Beyond the Covid dummy, the empirical design does not model multiple structural breaks, and merged estimates may mask regime-specific effects. External validity is therefore constrained to similar institutional contexts and to the 2012–2024 environment. Extending to other central banks could reveal different patterns driven by institutional design, credibility, and market structure, but such expansion lies beyond the scope of this thesis.

6. Conclusion

6.1 Summary

This thesis asked whether forward guidance has become more Odyssean , using more commitment in their policies, over 2012–2024. The evidence does not support an apparent shift in that direction. Across the Fed and the Riksbank institutions, Odyssean and Delphic elements coexist, with Delphic-classified language remaining consistently more prevalent in the measured distribution of words within forward-guidance speeches. This prevalence is best understood as a product of the measurement approach : under the dictionary-based approach adopted here, a larger share of tokens matches Delphic (forecast-oriented) terms than Odyssean (commitment-oriented) terms. It does not justify the stronger claim that central banks « conduct mainly Delphic guidance ». Rather, it indicates that the wording patterns captured by the dictionary classify a greater proportion of forward-guidance speech content as Delphic. Forward guidance is therefore best characterized, within this measurement framework, as a hybrid communication tool rather than one dominated by binding commitments (Moessner et al. 2014), consistent with the view that Odyssean and Delphic signals are distinct dimensions that can jointly appear in the same communication.

To address a significant gap in the literature, specifically the absence of a transparent quantitative measure of the degree of “Odysseanness” inherent in communication itself, the thesis developed lexicon-based tone indices that separately quantify Odyssean and Delphic content within each speech, complemented by negative and uncertainty controls. This measurement strategy avoids imposing a pre-determined binary classification of guidance and instead captures the diversity of tones within central-bank communication. The empirical design then assessed whether tone variation is reflected in market movements using OLS regressions on forward-guidance speech days, with absolute daily yield changes as the outcome. The 2-year yield served as the primary maturity given its closer link to expectations of the short-term policy path, while the 5-year yield provided a baseline maturity comparison.

The regression evidence indicates that, for the Fed, speech-level Odyssean and Delphic tone is not robustly associated with the magnitude of daily yield movements across baseline and robustness specifications. Instead, the most stable associations relate to contextual factors such as speaker authority and the Covid period, and effects are generally more pronounced at the shorter maturity than at the longer maturity. For the Riksbank, coefficient patterns are suggestive but cannot support strong inference due to limited sample size and weak model significance level. The Swedish results should therefore be interpreted as exploratory comparative evidence rather than a symmetric test. This pattern lends empirical support to recent policy critiques questioning the efficacy of explicit Odyssean guidance (English & Sack, 2024; Romer & Romer, 2024), particularly in light of the Fed’s post-2020 new policy framework , which markets did not fully incorporate before the subsequent policy shift in 2022.

Overall, the thesis develops a replicable two-dimensional framework for quantifying forward-guidance tone and finds that, with daily outcomes and dictionary-based measures , tone variation does not indicate a clear shift toward a more Odyssean style or a robust effect on Fed yield volatility. Limitations stem from the inherent constraints of lexicon methods (context, conditionality, evolving phrasing), sample strong difference across banks, and the noisiness of daily yield changes. Ultimately, the results imply that forward-guidance effectiveness is driven less by how strongly speeches sound « committed » and more by the credibility and coherence of the overall communication framework.

6.2 Recommendations for Future Research

The primary limitation of the current dictionary-based approach lies in its incapacity to capture tone nuance and syntactic dependency. Although transparent, word-count methods handle conditional phrasing, negation, and central-bank specific language poorly, making it difficult to distinguish firm commitments from conditional guidance
A necessary evolution would involve classifying communication at the sentence or paragraph level using transformer-based language models (e.g., BERT or specialized FinBERT, not currently existing). Unlike static lexicons, these models utilize contextual word embedding to distinguish between "we will not commit" and "we commit," capturing the intent behind bank-specific expressions. Future studies should use supervised machine learning by manually labeling a subset of guidance selections to train these models, which would reduce measurement error and limit attenuation bias in the estimated effects.

The use of daily absolute yield changes remains a noisy proxy for market reactions, as daily windows inevitably incorporate macroeconomic shocks and unrelated news cycles. To isolate the "pure" communication effect, future research should transition toward high-frequency event-study identification.By using narrow observation windows (like 15 to 30 minutes) around specific communications, researchers can more accurately attribute market volatility to the speech itself. Furthermore, decomposing the yield curve into policy-path expectations and term-premium components would provide deeper insight into the transmission mechanism. Such a decomposition would show whether forward guidance mainly affects expected future policy rates or term premium (risk compensation), two effects that are mixed together in daily yield data.

The focus on the Riksbank only limits the generalizability of the findings due to the specific institutional characteristics of the Swedish regime. To enhance external validity, research should expand the scope into a multi-banks panel dataset, incorporating central banks with varying communication mandates (e.g., the ECB, the Bank of England,…).Another point which could be interesting to analyze would be to expand the corpus to include minutes, press conferences, and reports. It would allow for a formal test of whether "tone effects" are consistent across different communication channels.

7. Appendices

Appendix 1 – Reporting the use of artificial intelligence (AI) during the thesis

1. AI applications used :

- Gemini (Google): Used as a linguistic assistant. Its role was limited to checking spelling mistakes , refining the academic tone, and ensuring the structural coherence of the author's original work.

- Perplexity AI & Connected Papers: These platforms were used for the identification and mapping of scholarly literature. Perplexity served as an exploratory search engine to locate relevant primary sources, while Connected Papers helped visualize the citation network and identify innovative works within the field.

- ChatGPT (OpenAI): Used as a computational and logic-based assistant. It served primarily as a support for technical « coding », specifically regarding Python scripting and the structural organization of analytical insights.

- Zotero: Used as a reference management system to ensure the systematized organization of the bibliography and adherence to citation standards (APA 7th edition).

2. Purpose of the AI used across the thesis structure

Throughout the drafting of the introduction, literature review, methodology, results, discussion, and conclusion, Gemini was used to refine already-written text. This intervention was strictly limited to improving clarity, ensuring a formal academic register, and correcting grammatical errors. The tool was never used to generate original arguments or synthesize findings.

For the literature review, Perplexity AI and Connected Papers were used to supplement traditional database searches (like Google Scholar). These tools facilitated the discovery of recent academic articles and helped outline the relationships between different theoretical frameworks. All sources identified through AI were independently read and verified by the author. Zotero was used to maintain bibliographic integrity and manage the technical aspects of citation.

For the empirical phase presented in the methodology, ChatGPT was used to develop and "debug" Python scripts executed via Google Colab. This involved an repeated process of trial and error to optimize data processing, implement complex econometric models, and generate codes for graphs and tables generation into Overleaf. While ChatGPT assisted in resolving syntax and clarifying coding logic, all data cleaning, model specifications, and statistical interpretations were directed and validated by the author. Finally, ChatGPT provided structural support during the results analysis phase by helping to organize and categorize the author's original observations and interpretations into a logical flow.

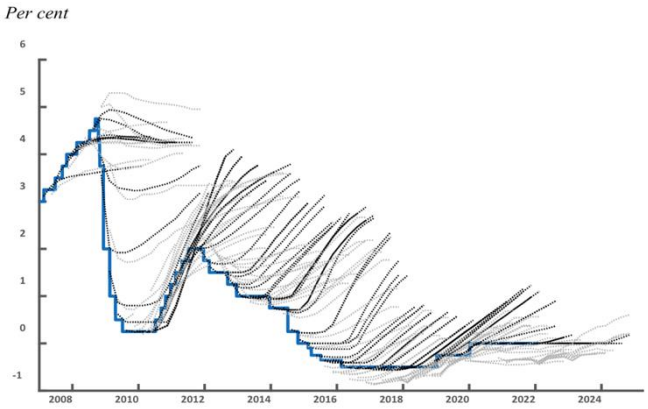
3. Ensuring Academic Integrity

- Verification of Sources: No citation provided by an AI tool was accepted at face value. Every reference was manually checked against original PDFs and university-provided databases to prevent the inclusion of "hallucinated" or inaccurate citations.

- Manual Data Analysis: All statistical outputs, econometric results, and data interpretations were produced through the author's work. AI was used to assist with the code that ran the analysis, but never to invent or alter the empirical results.

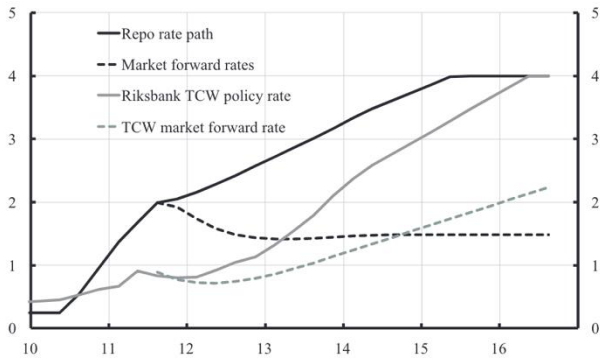
- Ethical Considerations: The use of AI in this thesis was guided by the principles of transparency and academic honesty. At no point was any AI tool used to generate original data, falsify empirical findings, or replace the fundamental tasks of research and writing. All AI-assisted content was supervised, edited, and contextualized by the author in accordance with the ethical guidelines for AI in higher education (Uliege Charter). The final document remains a reflection of the author's independent research and original analysis.

Appendix 2 – Riksbank's repo rate forecasts 2007-2021 (Flug & Honohan, 2022)

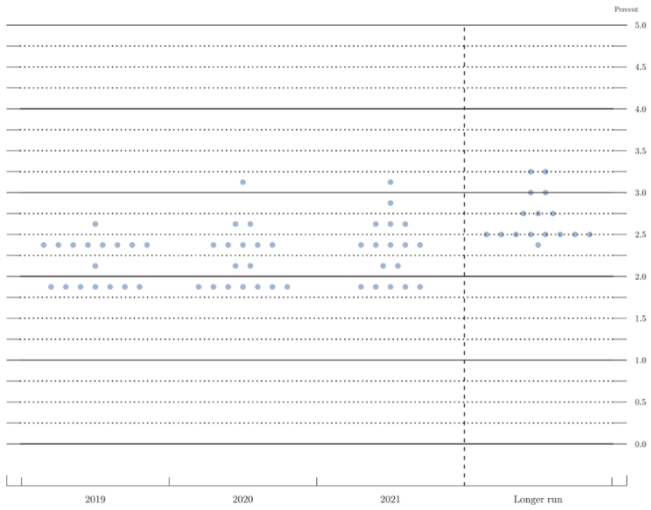


Note: The blue line is the repo rate, the black dotted lines are forecasts by the Riksbank and the grey dotted lines are market forecasts as quarterly averages at announcement dates. The market forecasts, as well as Stibor, are more volatile around the turn of the year since 2016, when the resolution fee was introduced.

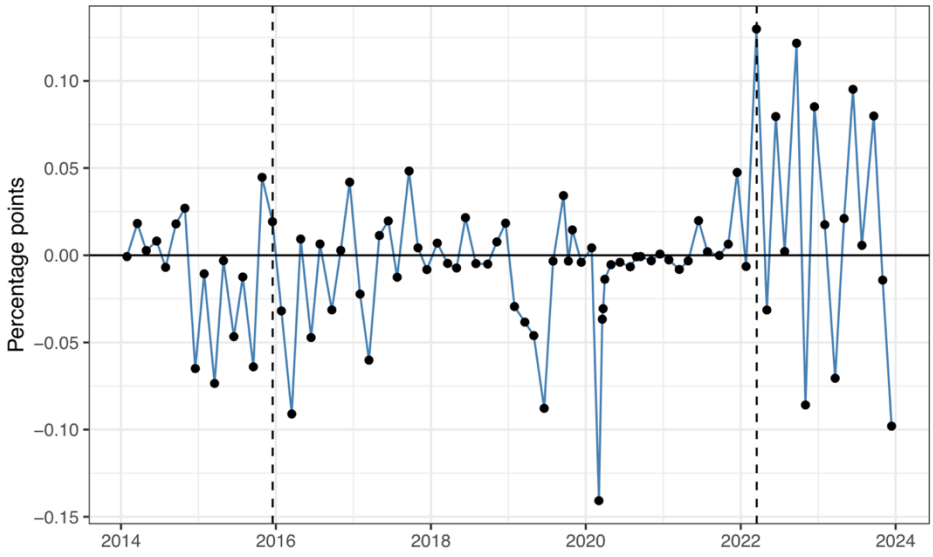
Appendix 3 – Riksbank and Market Policy Rate Paths, 2011 (Svensson, 2015)



Appendix 4 – June 2019 FED's Dot-plot (Bernanke, 2025).



Appendix 5 – Monetary Policy Surprises (Bauer et al., 2024)



Appendix 7 – VIF for the Fed and the Riksbank

Variance inflation factors (VIF) for Sveriges Riksbank

Regressor	VIF
Uncertainty_{Std}	3.642
Delphic_{Std}	3.215
Odyssean_{Std}	1.618
Negative_{Std}	1.482
Covid	1.432
Head	1.375
log(GapDays)	1.145

Variance inflation factors (VIF) for Federal Reserve

Regressor	VIF
Uncertainty_{Std}	2.299
Delphic_{Std}	2.249
Odyssean_{Std}	1.262
Negative_{Std}	1.184
Head	1.124
Covid	1.096
log(GapDays)	1.052

8. Bibliography

Aksit, D. (2020). Unconventional Monetary Policy Surprises : Delphic or Odyssean ? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3602291

Andersson, M., & Hofmann, B. (2009). Gauging the Effectiveness of Quantitative Forward Guidance : Evidence from Three Inflation Targeters. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1481577

Andersson, M., & Viotti, S. (1999). *Managing and preventing fínancial crises -lessons from the Swedish experience*. https://elischolar.library.yale.edu/cgi/viewcontent.cgi?article=11165&context=ypfs-documents

Andrade, P., & Ferroni, F. (2021). Delphic and odyssean monetary policy shocks : Evidence from the euro area. *Journal of Monetary Economics*, *117*, 816-832. https://doi.org/10.1016/j.jmoneco.2020.06.002

Apel, M., & Grimaldi, M. (2012). The information content of central bank minutes. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2092575

Assenmacher, K., Glöckler, G., Holton, S., Trautmann, P., Ioannou, D., Mee, S., Bakk-Simon, K., Bergbauer, S., Catenaro, M., Charalampakis, E., Ehrmann, M., Ferrero, G., Georgarakos, D., Gertler, P., Giovannini, A., Grandia, R., Hernborg, N., Herrala, N., Kedan, D., … Taylor, E. (2021). Clear, Consistent and Engaging : ECB Monetary Policy Communication in a Changing World. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3928296

Bassetto, M. (2019). *Forward Guidance : Communication, Commitment, or Both?* Federal Reserve Bank of Chicago. https://doi.org/10.21033/wp-2019-05

Bauer, M. D., Pflueger, C. E., & Sunderam, A. (2024). *Changing Perceptions and Post-Pandemic Monetary Policy*. https://www.kansascityfed.org/documents/10337/pflueger_jh.pdf

Benchimol, J., Kazinnik, S., & Saadon, Y. (2025). Federal Reserve Communication and the Covid -19 Pandemic. *The Manchester School*, *93*(5), 464-484. https://doi.org/10.1111/manc.12520

Bernanke, B. S. (2025). Improving Fed communications: a proposal', Hutchins Center Working Paper No. 102, Brookings Institution. https://www.brookings.edu/wp-content/uploads/2025/05/Hutchins-WP102.pdf

Blinder, A. S., Ehrmann, M., De Haan, J., & Jansen, D.-J. (2024). Central Bank Communication with the General Public : Promise or False Hope? *Journal of Economic Literature*, *62*(2), 425-457. https://doi.org/10.1257/jel.20231683

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D. (2008). Central Bank Communication and Monetary Policy : A Survey of Theory and Evidence. Journal Of Economic Literature, 46(4), 910-945. https://doi.org/10.1257/jel.46.4.910

Board of Governors of the Federal Reserve System. *The Fed Explained : What the Central Bank Does* (Eleventh). (2021). https://doi.org/10.17016/0199-9729.11

Breman, A., & Seim, A. (2025). *Openness and clarity – key ingredients in Riksbank communication*. *2*. https://www.riksbank.se/globalassets/media/rapporter/pov/artiklar/engelska/2025/250910/2025_2-openness-and-clarity--key-ingredients-in-riksbank-communication.pdf

Campbell, J. R., Evans, C. L., Fisher, J. D. M., & Justiniano, A. (2012). Macroeconomic Effects of Federal Reserve Forward Guidance. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2166310

Campbell, J. R., Fisher, J. D. M., Justiniano, A., & Melosi, L. (2017). Forward Guidance and Macroeconomic Outcomes since the Financial Crisis. *NBER Macroeconomics Annual*, *31*(1), 283-357. https://doi.org/10.1086/690242

Cieslak, A., & Schrimpf, A. (2019). Non-monetary news in central bank communication. Journal of International Economics, 118, 293–315. https://doi.org/10.1016/j.jinteco.2019.01.012

Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E., & Strasser, G. (2017). Communication of Monetary Policy in Unconventional Times. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3043098

Correa, R., Garud, K., Londono, J. M., & Mislang, N. (2017). Sentiment in Central Bank's Financial Stability Reports. International Finance Discussion Papers, 2017.0(1203), 1-46. https://doi.org/10.17016/ifdp.2017.1203

Del Negro, M., Giannoni, M. P., & Patterson, C. (2012). The Forward Guidance Puzzle. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2163750

Eggertsson, G. B., & Michael, P. W. (2003). Zero Bound on Interest Rates and Optimal Monetary Policy. Brookings Papers On Economic Activity, 2003(1), 139-233. https://doi.org/10.1353/eca.2003.0010

English, W. B., & Sack, B. (2024). Symposium on the Federal Reserve's Monetary Policy Framework Review. *Brookings Papers on Economic Activity*, *2024*(2), 3-92. https://doi.org/10.1353/eca.2024.a964369

Flug, K. (2022). *Evaluation of the Riksbank's monetary policy 2015-2020*. Sveriges riksdag. https://www.riksbank.se/globalassets/media/nyheter--pressmeddelanden/nyheter/2022/evaluation-of-the-riksbanks-monetary-policy-20152020.pdf

Goy, G., Hommes, C. H., & Mavromatis, K. (2018). Forward Guidance and the Role of Central Bank Credibility under Heterogeneous Beliefs. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3296214

Hagedorn, M., Luo, J., Manovskii, I., & Mitman, K. (2019). Forward guidance. *Journal of Monetary Economics*, *102*, 1-23. https://doi.org/10.1016/j.jmoneco.2019.01.014

Hansen, S., & McMahon, M. (2016). Shocking language : Understanding the macroeconomic effects of central bank communication. Journal of International Economics, 99, S114-S133. https://doi.org/10.1016/j.jinteco.2015.12.008

Hansson, M. (2021). Evolution of topics in central bank speech communication. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2109.10058

Hardeniya, T., & Borikar, D. A. (2016). Dictionary based approach to sentiment analysis: A review. International Journal of Advanced Engineering, Management and Science, 2(5), 317–322. https://ijaems.com/detail/dictionary-based-approach-to-sentiment-analysis-a-review/

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*(4), 709–722. https://doi.org/10.3758/bf03192961

Hubert, P., & Fabien, L. (2017). Central Bank sentiment and policy expectations. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2920496

Istrefi, K., Odendahl, F., & Sestieri, G. (2023). Fed communication on financial stability concerns and monetary policy decisions: Revelations from speeches. Journal of Banking & Finance, 151, 106820. https://doi.org/10.1016/j.jbankfin.2023.106820

Labonte, M. (2021). The Federal Reserve's Response to Covid -19 : Policy Issues. *The Federal Reserve*. https://elischolar.library.yale.edu/cgi/viewcontent.cgi?article=1199&context=ypfs-documents2

Levin, A., & Sinha, A. (2020). *Limitations on the Effectiveness of Monetary Policy Forward Guidance in the Context of the* Covid *-19 Pandemic* (No. w27748; p. w27748). National Bureau of Economic Research. https://doi.org/10.3386/w27748

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & McDonald, B. (2012). IPO First-Day returns, offer price revisions, volatility, and Form S-1 language. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2128766

Lu, X., & White, H. (2013). Robustness checks and robustness tests in applied economics. *Journal of Econometrics*, *178*, 194–206. https://doi.org/10.1016/j.jeconom.2013.08.016

MacKinnon, J. G., & White, H. (1985b). Some heteroskedasticity-consistent covariance matrix estimators
with improved finite sample properties. Journal of Econometrics, 29(3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7

Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. Journal of King Saud University - Computer and Information Sciences, 36(4), 102048. https://doi.org/10.1016/j.jksuci.2024.102048

Masciandaro, D., Peia, O., & Romelli, D. (2024). Central bank communication and social media : From silence to Twitter. *Journal of Economic Surveys*, *38*(2), 365-388. https://doi.org/10.1111/joes.12550

McKay, A., Nakamura, E., & Steinsson, J. (2016). The Power of Forward Guidance Revisited. *American Economic Review*, *106*(10), 3133-3158. https://doi.org/10.1257/aer.20150063

Mester, L. J. (2024). *Forward Guidance and Monetary Policy Communications : Use Your Words and Connect the Dots*. https://www.clevelandfed.org/collections/speeches/2024/sp-20240528-forward-guidance-and-monetary-policy-communications
Moessner, R., Haan, J. <., & Jansen, D. (2014). The Effect of the Zero Lower Bound, Forward Guidance and Unconventional Monetary Policy on Interest Rate Sensitivity to Economic News in Sweden. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2386705

Moessner, R., Jansen, D., & De Haan, J. (2015). Communication About Future Policy Rates in Theory and Practice : A Survey. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2616605

Nakamura, E., & Steinsson, J. (2013). High Frequency Identification of Monetary Non-Neutrality : the Information Effect. The Quarterly Journal Of Economics, 133(3), 1283-1330. https://doi.org/10.3386/w19260

Natvik, G. J., Rime, D., & Syrstad, O. (2020). Does publication of interest rate paths provide guidance? *Journal of International Money and Finance*, *103*, 102123. https://doi.org/10.1016/j.jimonfin.2019.102123

Neely, C. J. (2014). Unconventional monetary policy had large international effects. Journal Of Banking & Finance, 52, 101-111. https://doi.org/10.1016/j.jbankfin.2014.11.019

Papell, D. H., & Prodan-Boul, R. (2024). Policy rules and forward guidance following the Covid-19 recession. Journal Of Financial Stability, 74, 101321. https://doi.org/10.1016/j.jfs.2024.101321

Picault, M., & Renault, T. (2017). Words are Not All Created Equal: A New Measure of ECB Communication. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2980777

Romer, C. D., & Romer, D. H. (2024). Did the Federal Reserve's 2020 Policy Framework Limit Its Response to Inflation ? Evidence and Implications for the Framework Review. Brookings Papers On Economic Activity, 2024(2), 59-87. https://doi.org/10.1353/eca.2024.a966123

Siino, M., Tinnirello, I., & La Cascia, M. (2023). Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers. *Information Systems*, *121*, 102342. https://doi.org/10.1016/j.is.2023.102342

Singh, A. (2023). The Role of Central Banks in Managing Monetary Policy and its Effects on the Economy : An Empirical Analysis of Policy Tools, Economic Indicators, and Implications. *Journal of Student Research*, *12*(4). https://doi.org/10.47611/jsrhs.v12i4.5327

Statement on Longer-Run Goals and Monetary Policy Strategy. (2025). https://www.federalreserve.gov/monetarypolicy/files/FOMC_LongerRunGoals.pdf

Svensson, L. E. O. (2015). Forward Guidance. *International Journal of Central Banking*, *11*. https://larseosvensson.se/files/papers/forward-guidance.pdf

Swanson, E. (2017). Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets. https://doi.org/10.3386/w23311

Swanson, E. T. (2021). Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, *118*, 32-53. https://doi.org/10.1016/j.jmoneco.2020.09.003

Swanson, E. T., & Jayawickrema, V. (2024). *Speeches by the Fed Chair Are More Important Than FOMC Announcements : An Improved High-Frequency Measure of U.S. Monetary Policy Shocks*. https://sites.socsci.uci.edu/~swanson2/papers/hfdat.pdf

Université de Liège. (2023, December 22). ULiège Charter for the use of generative artificial intelligence in academic work. Retrieved January 6, 2026,

from https://www.student.uliege.be/cms/c_19230399/en/uliege-charter-for-the-use-of-generative-artificial-intelligence-in-academic-work

Vijayarani, S., & Janani, R. (2016).Text Mining: open Source Tokenization Tools – An Analysis. Advanced Computational Intelligence an International Journal (ACII), 3(1), 37–47. https://doi.org/10.5121/acii.2016.3104

Woodford, M. (2012). Methods of Policy Accommodation at the Interest-Rate Lower Bound. 2012 Jackson Hole Economic Policy Symposium, 185-288. https://doi.org/10.7916/d8z899cj

Wooldridge, J. M. (2009). Introductory econometrics: A modern approach (4th ed., student ed.). South-Western.

Xu, X., & Xu, X. (2023). Forward guidance investigation in new Keynesian models. *Economic Research-Ekonomska Istraživanja*, *36*(2), 2145983. https://doi.org/10.1080/1331677X.2022.2145983

EXECUTIVE SUMMARY

This thesis investigates whether central bank forward guidance has evolved into a more "Odyssean" (commitment-based) form over the last decade and evaluates its impact on financial market stability. The study is motivated by the tension between theoretical arguments for binding commitments at the effective/zero lower bound and the practical preference of policymakers for "Delphic" flexibility. To explore this, the research conducts a comparative analysis of the Federal Reserve and Sveriges Riksbank, examining their communication strategies across varied macroeconomic regimes, including the Covid-19 pandemic.

Methodologically, the study employs a dictionary-based sentiment analysis using a customized version of the Loughran and McDonald financial lexicon, extended with specific terms derived from Campbell et al. (2012) to distinguish between Odyssean and Delphic tones. These sentiment indices are tested against daily absolute changes in 2-year and 5-year government bond yields using an OLS regression framework with robust standard errors.

The results indicate that for the Federal Reserve, neither Odyssean nor Delphic tones is robustly associated with yield volatility in baseline specifications. During the Covid-19 crisis, higher Odyssean tone was associated with lower volatility in 2-year yields in some specifications, consistent with a potentially stronger role for commitment wording under extreme uncertainty. Exploratory findings for the Riksbank suggest a positive association between commitment language and volatility, though these results are constrained by sample size limitations.

By integrating natural language processing with financial econometrics, this research contributes a transparent design to quantify forward guidance types and highlights the empirical challenges of mapping theoretical policy constructs into textual indicators, suggesting that communication effects are context-dependent rather than structurally uniform.

KEYWORDS: FED – Riksbank – Central bank communication – Speeches – Sentiment analysis – Dictionary-based approach – Absolute yield change – Monetary policy – Forward Guidance – Zero/Effective Lower Bound – NLP

WORD COUNT: 18.295

Ecole de Gestion de l'Université de Liège