Article

"Decoding" Policy Perspectives: Structural Topic Modeling of European Central Bankers' Speeches

Thierry Warin [1] and Aleksandar Stojkov [2,*]

[1] Department of International Business, HEC Montreal, Montréal, QC H3T 2A7, Canada; thierry.warin@hec.ca
[2] Department of Business Law and Economics, Iustinianus Primus Law Faculty, Ss. Cyril and Methodius University, 1000 Skopje, North Macedonia
[*] Correspondence: a.stojkov@pf.ukim.edu.mk

Abstract: This research analyzes the speeches and interviews of high-level European Central Bank decision makers from 1997 to 2021 to identify clusters of prominent topics. Transparency is a crucial aspect of modern monetary policy, and public speeches and interviews by central bankers play a vital role in achieving it. Our study employs structural topic modeling to compare the prevalence of topics during the Global Financial Crisis, pandemic crisis, and periods of heightened inflation. Additionally, we explore the impact of central bank independence on official rhetoric.

1. Introduction

Public speeches and interviews delivered by high-ranking officials are essential communication tools for conveying the policy decisions and intentions of modern central banks. In this study, we aim to uncover clusters of prominent topics contained in the speeches and interviews delivered by central bankers participating in key decision-making bodies of the European Central Bank (ECB) during the 1997–2021 period. These speeches are widely regarded as aggregators of quantitative and qualitative analysis by supreme monetary institutions concerning macroeconomic developments, the monetary policy stance, and the soundness of the financial system (e.g., Petropoulos and Siakoulis 2021).

Our central research objective is to identify top topics in European central bankers' speeches and interviews, assess correlations among topics, and formulate policy recommendations for designing more effective forward guidance and enhancing the credibility of monetary policy. More specifically, our research goals are to: (1) identify topic prevalence and content during the 2008 Global Financial Crisis 2008, the SARS-CoV II pandemic crisis (2020–2021), and various periods of heightened inflation; (2) investigate whether there is a systematic difference in topic prevalence between more independent and less independent central banks; and (3) draw policy-relevant conclusions for designing forward guidance and enhancing the credibility of central banks.

To achieve these goals, we employ a text mining technique, specifically structural topic modeling, to uncover clusters of main topics in central bankers' speeches during the 1997–2021 period and various sub-periods of interest. Contrary to the naïve expectation that monetary policy tightening will be a widely exploited topic during periods of higher inflation, structural topic modeling offers a more systematic and nuanced assessment of the central topics covered, identification of clusters of topics, and their relative importance during different sub-periods.

In sum, this study contributes to the rapidly growing empirical literature on central bank communication and transparency by combining quantification through central bank transparency indices and text mining techniques to identify policy concerns, decisions,

and intentions. Our analysis highlights the value of these methods in understanding central bank communication and transparency, offering valuable insights for policymakers, financial markets, and researchers.

Subsequently, we will commence with the literature review, followed by outlining the empirical strategy and data collection. The model calibration, results, and conclusion will be presented thereafter.

2. Literature Review

Central bank transparency was initially a vague concept, with Issing (2005) broadly defining it as an explanation of monetary policy decisions to the public. Geraats (2002, p. 533) later conceptualized transparency as "the absence of asymmetric information between monetary policymakers and other economic agents", emphasizing the reduction of uncertainty as the primary benefit of transparency. According to the ECB, transparency implies that "the central bank provides the general public and the markets with all relevant information on its strategy, assessments, and policy decisions as well as its procedures in an open, clear, and timely manner".

Empirical research (e.g., Geraats 2002; Eijffinger and Geraats 2006; Dincer et al. 2019) introduced a richer conceptual framework, encompassing several dimensions of central bank transparency: political, economic, procedural, policy, and operational transparency (a thorough theoretical review is provided by Geraats 2002). These aspects have been codified and monitored across countries over time.

Political transparency involves openness about policy objectives and institutional arrangements, clarifying the motives of monetary policymakers. This includes elements such as explicit inflation targets, central bank independence, and contracts. Economic transparency focuses on publishing economic information relevant to monetary policy decisions, including economic data, policy models, and central bank forecasts. Procedural transparency explains how monetary policy decisions are made, providing a description of the strategy and an account of policy deliberations through minutes and individual voting records. Policy transparency entails promptly announcing and explaining policy decisions and providing indications of likely future policy actions. Operational transparency is related to the implementation of monetary policy actions and addresses potential complications arising from control errors or unanticipated macroeconomic disturbances.

Dincer et al. (2022) show that the ECB's transparency has considerably improved over time, particularly in economic and policy transparency (Figure 1). In this context, we argue that the Natural Language Processing (NLP) of public speeches and interviews of high-ranking ECB officials can provide valuable insights into the bank's progress in enhancing transparency.

Central bank transparency has become an indispensable feature of modern monetary policy, with public speeches and interviews serving as essential components. Transparency imposes self-discipline on central banks, enhances credibility by anchoring inflation expectations, and increases predictability of monetary policy outcomes and actions. The art of managing private expectations has been recognized as a central feature of monetary theory (e.g., Morris and Shin 2002; Woodford 2005). The global trend towards greater central bank independence is widely considered a strong driving force of increased central bank transparency (e.g., Garriga 2016).

Our study connects two broad strands of empirical literature devoted to central bank communication and transparency. The first strand focuses on quantification through central bank or monetary policy transparency indices, measuring the degree of information disclosure (e.g., Geraats 2002, 2020; Dincer and Eichengreen 2014; Dincer et al. 2019, 2022). The second strand employs text mining techniques to identify policy concerns, decisions, and intentions (e.g., Masawi et al. 2014; Hansen and McMahon 2016; Hansen et al. 2018, 2019; Warin and Sanger 2020; Küsters 2022). Our contribution relates to the second strand of rapidly growing empirical literature exploring central bank communication from a computational linguistics perspective.
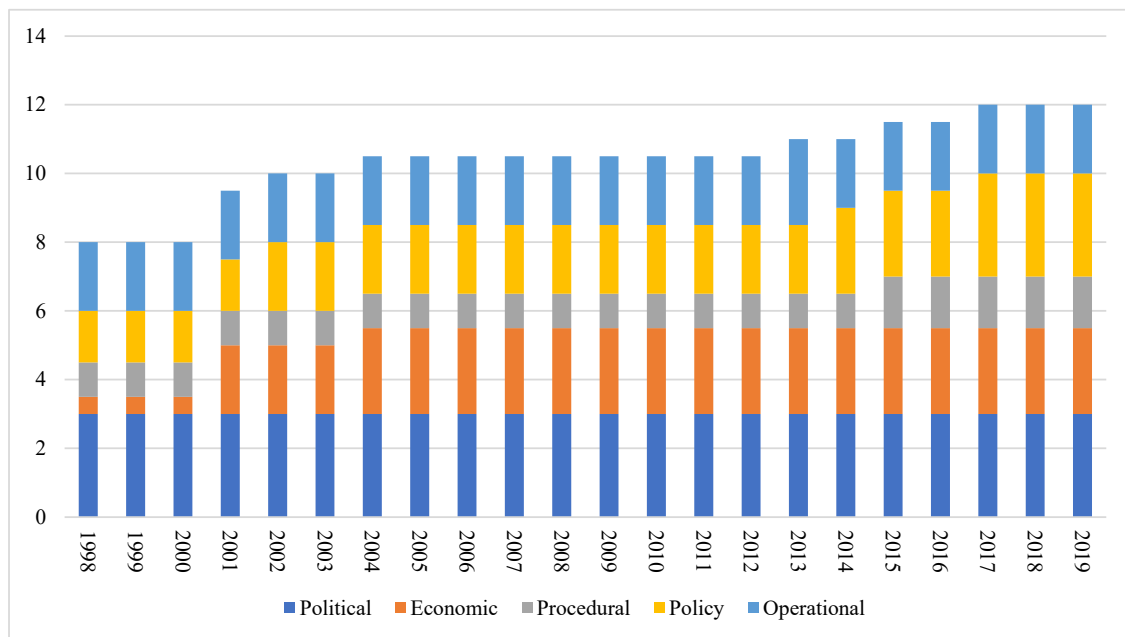
Figure 1. Components of the transparency index of ECB (1998–2019). Source: Based on data from
Dincer et al. (2022).

Central bank communication is critically important for the effectiveness of monetary
policy. There is a voluminous and constantly growing body of literature on the topic (e.g.,
Coibion et al. 2019, 2022; Cortes et al. 2022, 2023; Gorodnichenko et al. 2023). In this context,
Cortes et al. (2023) investigate the communication by the Federal Reserves, underscoring
that it plays a global role and its monetary policy decisions affect other central banks and
economies (Dedola et al. 2020). This strand of the empirical literature investigates the
signaling channel of the monetary policy, which, in turn, is increasingly perceived as the
'art of managing expectations' of private agents.

To do so, text mining has been used to filter out the most important messages and
signals in key speeches and interviews in the central banking world. Official monetary
statements by central bankers have been studied in connection with stock market perfor-
mance (e.g., Bennani 2019; Mathur and Sengupta 2019; Sapphasak 2019; Maqsood et al.
2020; Petropoulos and Siakoulis 2021), short-term interest rates or money markets (e.g.,
Hubert and Labondance 2017; Su et al. 2019), policy interest rates (e.g., Apel et al. 2022),
crisis management (e.g., Kahveci and Odabaş 2016), forward guidance (e.g., Shapiro and
Wilson 2021), and the measurement of central bank credibility (e.g., Warin and Sanger 2020),
among others. More specifically, the study of Cortes et al. (2023) is particularly relevant in
that it shows the content similarity of what is discussed in central bank communications
and the manager-analyst dialogues in earnings conference calls.

3. Empirical Strategy and Data Collection

We propose the following empirical strategy to analyze the informational content
of public speeches and interviews delivered by ECB's high-ranking officials (presidents,
vice-presidents, and other members of the Executive Board and of the Governing Council).

Public speeches and interviews delivered by high-ranking officials are essential com-
munication tools for conveying the policy decisions and intentions of modern central banks.
In this study, we aim to uncover clusters of prominent topics contained in the speeches
and interviews delivered by central bankers participating in key decision-making bodies of
the European Central Bank (ECB) during the 1997–2021 period. These speeches are widely
regarded as aggregators of quantitative and qualitative analysis by supreme monetary
institutions concerning macroeconomic developments, the monetary policy stance, and the
soundness of the financial system (Petropoulos and Siakoulis 2021).

Our central research objective is to identify top topics in European central bankers' speeches and interviews, assess correlations among topics, and formulate policy recommendations for designing more effective forward guidance and enhancing the credibility of monetary policy. More specifically, our research goals are to: (1) identify topic prevalence and content during significant economic events, such as the Global Financial Crisis (Bernanke 2004) and the pandemic crisis (Summers 2020); (2) analyze the relationship between economic conditions and the prominence of certain topics in speeches; and (3) uncover the differences in the topics emphasized by different ECB officials, based on their roles, backgrounds, or countries' levels of central bank independence, as measured by Grilli et al. (1991).

To achieve these goals, we employ a text mining technique, specifically structural topic modeling (STM), to uncover clusters of main topics in central bankers' speeches during the 1997–2021 period and various sub-periods of interest. We collected 2491 speeches and interviews and their metadata (date, speakers, title, and subtitles) from 7 February 1997 to 20 October 2021. This period covers much of the turbulent transition from the Great Moderation (Bernanke 2004) to the Secular Stagnation (Summers 2020). The data are gathered using an R-based program that interrogated the streaming Application Programming Interface (API) of the central bankers' speeches database maintained by the Bank for International Settlements.

We preprocess the data by cleaning and tokenizing the text, removing stop words, and stemming or lemmatizing the words to reduce dimensionality. We then select an appropriate STM model and input parameters, such as the number of topics, prior distributions, and model assumptions. By integrating metadata, we can analyze the relationship between metadata and topics, identify trends, and uncover patterns in the speeches. We extract and interpret the identified topics by examining the most representative words for each topic, as well as analyzing the relationships between topics and metadata. We visualize the results using various tools, such as heatmaps, network graphs, or word clouds.

The first findings indicate that European central bankers have become more vocal over time, especially during periods of economic turmoil (Figure 2). The largest number of speeches and interviews are concentrated around significant economic events, such as the introduction of the euro, the aftermath of the Global Financial Crisis (Bernanke 2004), and the 2017–2018 period.

As shown in Figure 3, November emerges as the month when central bankers frequently deliver important messages to the public, while August is generally considered a low season for such communications. The high prevalence of speeches and interviews in November can be attributed to the need for central bankers to explain expectations and deliberations for the upcoming calendar year, providing guidance to the public and financial markets. In terms of the days of the week, speeches and interviews are typically scheduled either before or after the weekend, with very few instances taking place during the weekend itself. Interestingly, Monday appears to be the most preferred day for central bankers to address the general audience or financial markets, potentially setting the tone for the week ahead.

Table 1 presents the speaker count proportions for various ECB officials, showcasing the number of speeches and interviews each official has delivered and their corresponding proportion of the total speeches in the dataset.

Jean-Claude Trichet leads the list with 329 speeches and interviews, making up 13.2% of the total. Benoît Coeuré follows with 190 speeches and interviews, constituting 7.6% of the total. Mario Draghi is not far behind with 186 speeches and interviews, accounting for 7.5% of the total. Other notable speakers include Yves Mersch (6.5%), Gertrude Tumpel-Gugerell (6.3%), and Willem F. Duisenberg (6.3%).

The table highlights the distribution of communication activities among ECB officials. It demonstrates that certain officials, such as Jean-Claude Trichet and Mario Draghi, have played a more prominent role in communicating the ECB's policy decisions and intentions to the public.

Figure 2. Number of speeches and interviews over time. Note: Speeches and interviews delivered by high ECB officials. Source: Authors' calculations based on data from the Bank for International Settlements (2023).

Figure 3. Number of speeches and interviews over time (1997–2021). Note: Speeches and interviews delivered by high-ranking ECB officials. Source: Authors' calculations based on data from the Bank for International Settlements (2023).

Table 1. Speaker count proportions.

Speaker	Number of Speeches and Interviews	Proportion of Total
Jean-Claude Trichet	329	13.2%
Benoît Coeuré	190	7.6%
Mario Draghi	186	7.5%
Yves Mersch	161	6.5%
Gertrude Tumpel-Gugerell	157	6.3%
Willem F. Duisenberg	157	6.3%
Peter Praet	126	5.1%
Vítor Constâncio	126	5.1%
José Manuel González-Páramo	119	4.8%
Lorenzo Bini Smaghi	119	4.8%
Otmar Issing	100	4.0%
Jürgen Stark	88	3.5%
Lucas Papademos	87	3.5%
Sabine Lautenschläger	84	3.4%
Eugenio Domingo Solans	76	3.1%

Jean Claude-Trichet holds the distinction of being the most communicative president of the ECB, as evident from Table 1. During Mario Draghi's tenure, the vice-presidents and members of the Executive Board also played an active role in communication, as illustrated in Figure 4. Christine Lagarde, being a relatively recent ECB president during the period under study, is underrepresented in the sample. As her term progresses, it is expected that her communication frequency will increase, potentially altering the overall distribution of central bankers' speeches and interviews.

Table 2 presents the number of speeches and interviews delivered by ECB officials from different countries, grouped into categories based on the level of central bank independence.

Table 2. Number of speeches and interviews per category.

Category	Country	Count Proportion
Category 1	Greece	87 (3.5%)
	Italy	454 (18.2%)
	Portugal	126 (5.1%)
	Spain	175 (7.0%)
Category 2	Austria	157 (6.3%)
	Belgium	133 (5.3%)
	France	611 (24.5%)
	Ireland	24 (10%)
Category 3	Finland	43 (1.7%)
	Germany	513 (20.6%)
	Luxembourg	161 (6.5%)
	Netherlands	7 (0.3%)

In Category 1, which represents countries with less independent central banks, Italy has the highest count proportion with 454 speeches and interviews, accounting for 18.2% of the total. This is followed by Spain (7.0%), Portugal (5.1%), and Greece (3.5%).

Category 2 includes countries with a moderate level of central bank independence. France has the highest count proportion in this group with 611 speeches and interviews, making up 24.5% of the total. Other countries in this category include Belgium (5.3%), Austria (6.3%), and Ireland (10%).

Lastly, Category 3 represents countries with the most independent central banks. Germany leads this group with 513 speeches and interviews, accounting for 20.6% of the total. Luxembourg contributes 6.5%, while Finland and the Netherlands have relatively lower proportions at 1.7% and 0.3%, respectively.

The table shows that countries with higher central bank independence (Category 3) tend to have a larger share of speeches and interviews. It also highlights the significant communication activities from countries like France, Italy, and Germany.
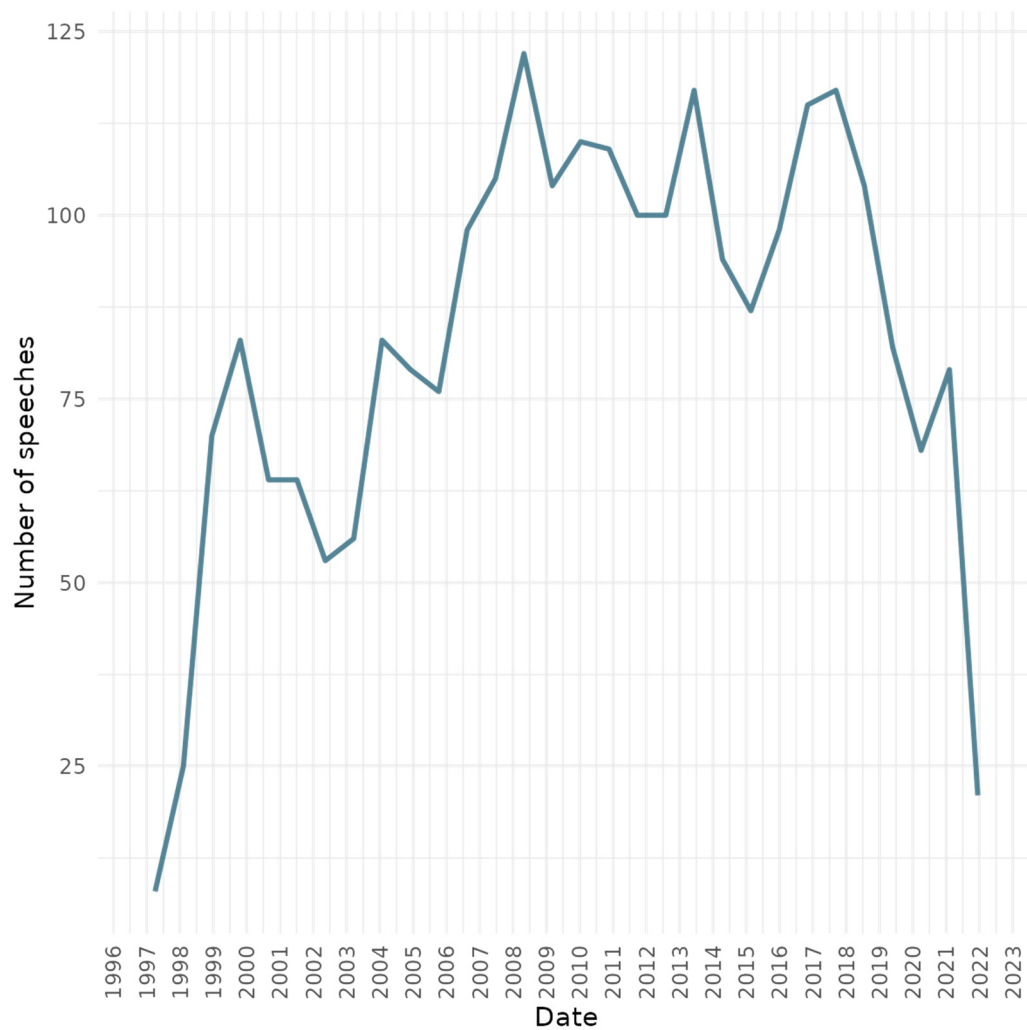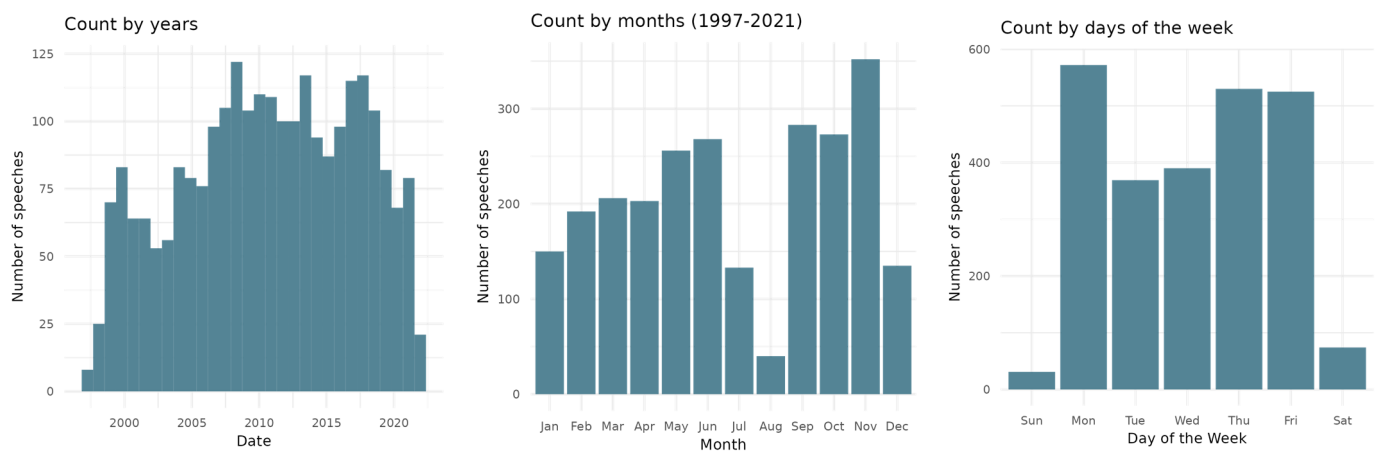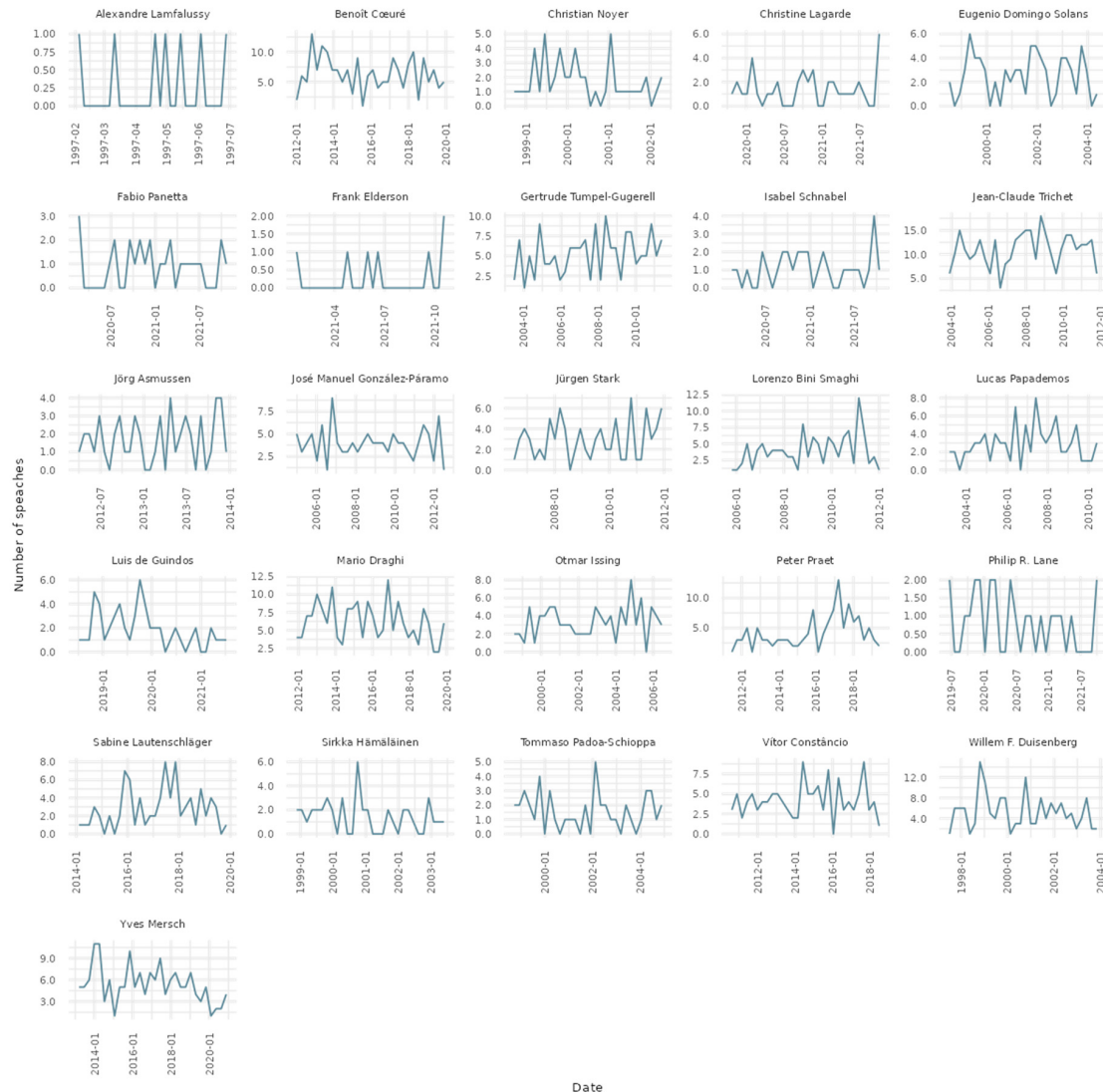

Figure 4. Number of speeches per speaker over time. Note: Speeches and interviews delivered by high ECB officials. Source: Authors' calculations based on data from the Bank for International Settlements (2023).

4. Model Selection and Calibration

In recent years, public speeches and interviews delivered by high-ranking officials have become a crucial means for central banks to communicate their policy decisions and intentions. Understanding the content of these communications is essential for anticipating policy directions and market expectations. To this end, we employ Structural Topic Modeling (STM) to analyze the speeches of European Central Bank (ECB) officials in order to uncover patterns, emerging themes, and their relation to economic conditions.

We compile a comprehensive dataset of speeches and interviews from high-profile European central bankers, sourced from Bank for International Settlements (2023). The data preprocessing involves cleaning the text, tokenizing, removing stop words, and stemming or lemmatizing words to reduce dimensionality. Moreover, we incorporate metadata, such as speaker information, dates, and economic indicators, to enhance the analysis.

STM, building on the tradition of probabilistic topic models like Latent Dirichlet Allocation and the Correlated Topic Model, is an innovative approach that incorporates arbitrary metadata into the topic model. This technique allows for summarizing unstructured text, discovering hidden topic clusters, and assigning probabilities of belonging to specific topics for each document. We use the stm package for our analysis.

Following a typical STM workflow, as illustrated in Figure 5, we estimate the STM after ingesting and preparing the data. Rapid estimation facilitates the evaluation, comprehension, and visualization of results. We select appropriate model parameters and iteratively refine the topic estimates until convergence. Once the identified topics are extracted, we interpret the most representative words for each topic and analyze the relationships between the topics and metadata. We also use various visualization tools, network graphs, or word clouds to display the results and explore patterns in central bankers' communication.
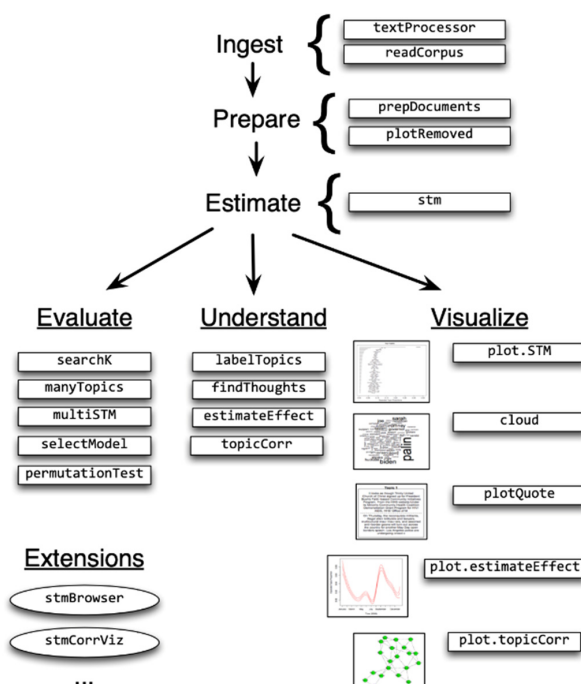

Figure 5. STM heuristic description.

Our STM-based analysis of high-profile European central bankers' speeches and interviews holds significant implications for policymakers, financial markets, and researchers. By understanding trends and patterns in communication, market participants can better anticipate policy decisions, and policymakers can assess the effectiveness of their communication strategies.

The workflow for Structural Topic Modeling is a streamlined process involving four key steps. Initially, the algorithms read and process the text data, preparing it for further analysis. In the second step, the text is combined with relevant metadata to enrich the model's understanding of the documents. Following this, the model estimates the topics present in the data and evaluates their statistical significance. Finally, the results are visualized, making them easier to conceptualize and interpret, thereby providing valuable insights into the underlying themes and patterns present in the text data.

4.1. Estimation with a Topical Prevalence Parameter

In this study, we incorporate the "category" variable (level of independence of a central bank in monetary policy) and the "day" variable (integer measure of days running from the first to the last day of a year) as covariates in the topic prevalence portion of the STM model applied to the ECB speeches and interviews data.

Topic prevalence captures the extent to which each topic contributes to a document, with each document being modeled as a mixture of multiple topics. Given that different documents come from various sources, it is reasonable to allow the prevalence to vary with metadata associated with document sources.

By incorporating the "category" variable, coded as "C1", "C2", or "C3", and the "day" variable into the topic prevalence estimation, we can analyze how the central bank's independence level and the time of year influence the topics discussed in speeches and interviews. This approach allows us to uncover patterns and trends related to the prominence of certain topics based on the central bank's independence level and the day of the year, providing valuable insights for policymakers, financial markets, and researchers.

4.2. Model Search across Numbers of Topics

In the STM analysis, selecting the appropriate number of topics for a given corpus is crucial, though there is no definitive "right" answer. To address this issue, the searchK function provides a data-driven approach to choosing the number of topics by performing several automated tests and calculating various quantities of interest.

The searchK function calculates the held-out likelihood (Wallach et al. 2009) and performs a residual analysis (Taddy 2015) to help determine the number of topics. Additionally, it computes quantities of interest, such as the average exclusivity and semantic coherence (Roberts et al. 2019), which can further inform the selection process.

In our study, we estimated 10 to 20 topics for each STM model and compared the results based on the aforementioned criteria. To select the optimal number of topics, we followed a two-step process: (1) retain models with the minimum residual: by focusing on models with the lowest residuals, we aimed to reduce the error in the topic estimation and improve the model's performance, and (2) choose the model with the maximum semantic coherence: from the models with minimum residual, we selected the one with the highest semantic coherence. This ensured that the topics generated by the STM model were meaningful and interpretable, which is crucial for effective analysis and communication of the results.

By following this approach, we were able to identify an optimal number of topics for our STM analysis that balanced the need for accurate topic estimation and meaningful interpretation of the results.

4.3. Topical Prevalence Contrast: Central Bank Independence

Visualizing metadata/topic relationships is a key advantage of the STM model, as it allows researchers to examine the connection between metadata variables and the estimated topics. In order to plot these relationships, users need to specify the variable of interest for calculating an effect.

The effect can be represented as the expected proportion of a document that belongs to a topic as a function of a covariate, or as a first difference type estimate, where topic prevalence for a specific topic is contrasted for two or more groups.

When the covariate of interest is binary or when users are interested in a particular contrast, the method = "difference" option is useful for plotting the change in topic proportion when shifting from one specific value to another (Roberts et al. 2019). In our analysis, we chose the independence of central banks in pursuing monetary policy as the covariate.

By employing the method = "difference" option, we can visualize how the topic prevalence varies between different levels of central bank independence. This enables us to better understand the relationship between central bank independence and the topics discussed in their speeches and interviews. As a result, we can gain valuable insights into the policy priorities and communication strategies of central banks with varying degrees of independence in the context of monetary policy.

5. Results

5.1. Entire Period (1997–2021)

In Figure 6 we can observe the diagnostic values for the different topic models we estimated, including the held-out likelihood and residual values for each number of topics. We also report the number of documents, words, and tokens removed by different thresholds during the preprocessing stage. This information helps assess the quality of the data and the appropriateness of the chosen model. For example, if the held-out likelihood does not increase significantly as we add more topics, it may suggest that the model is overfitting the data. Additionally, if the residual values are high, it may indicate that the model is not capturing all of the variation in the data.
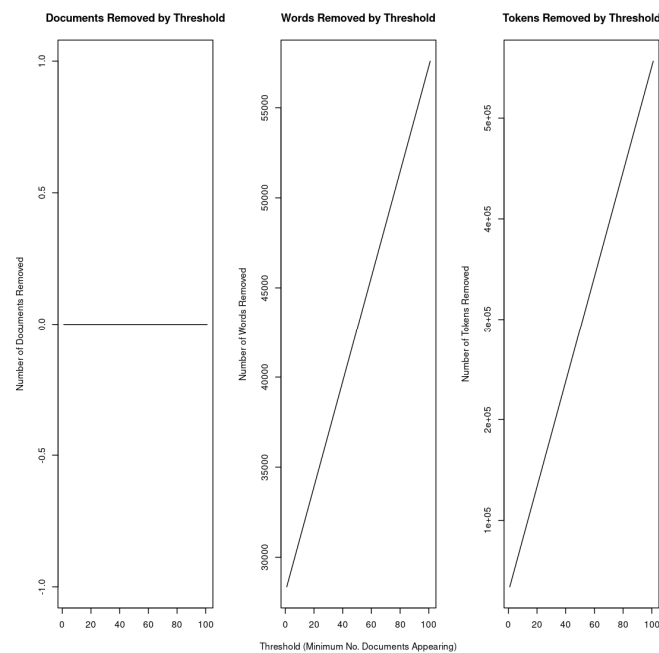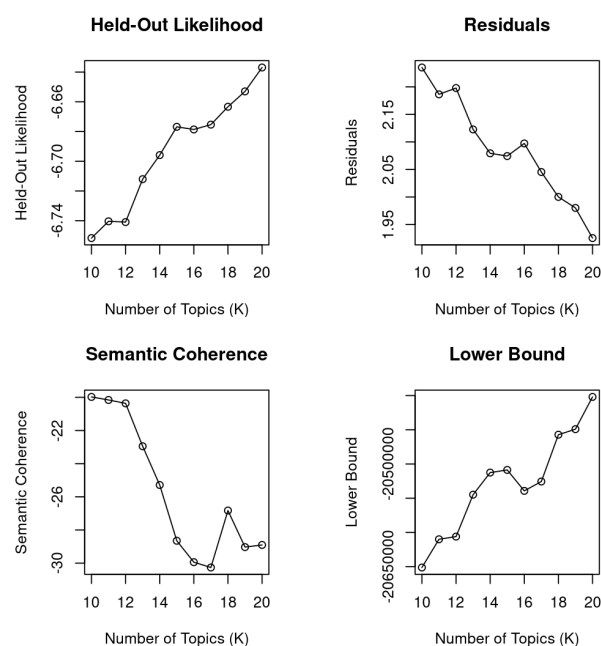
Figure 6. Number of words and documents removed for different thresholds (1997–2021).

Figure 7 presents the top topics discussed in the speeches and interviews of high-profile European Central Bank officials during the entire period under consideration (1997–2021). The figure shows the expected proportion of the corpus that belongs to each topic in Panel A. Panel B of the figure shows the correlations between the top topics.

Panel A. Top topics

Panel B. Correlations among them

Figure 7. Top topics (1997–2021).

The most used topics in the speeches and interviews during the entire period under consideration (1997–2021) were [T-18] monetary, policy, stability; [T-17] financial, bank, risk; [T-16] inflation, rate, policy; [T-9] euro, crisis, fiscal; and [T-12] growth, euro, countries. The corpus-level visualization is in terms of the expected proportion of the corpus that belongs to each topic (Figure 7, Panel A).

In terms of topic correlations, [T-18] monetary, policy, stability is intertwined with [T-13] market, bank, liquidity, and [T-14] rate, inflation, policies, which, as a larger cluster of topics, is also correlated with [T-6] bank, financial, capital (Figure 7, Panel B). These four topics are connected via [T-17] financial, bank, risk and [T-9] euro, crisis, fiscal, with [T-5] European, Union, Europe, and [T-18] monetary, policy, stability, cutting to the core of the ECB's mandate.

Figure 8 shows the word clouds of the top five topics discussed in the speeches and interviews of high-profile European central bankers during the period of 1997–2021. The size of each word represents its frequency within the topic. The top five topics are "monetary, policy, stability" (T-18), "financial, bank, risk" (T-17), "inflation, rate, policy" (T-16), "euro, crisis, fiscal" (T-9), and "growth, euro, countries" (T-12). The word clouds provide a visual representation of the most frequently used words in each topic, giving insights into the main themes discussed by the central bankers.

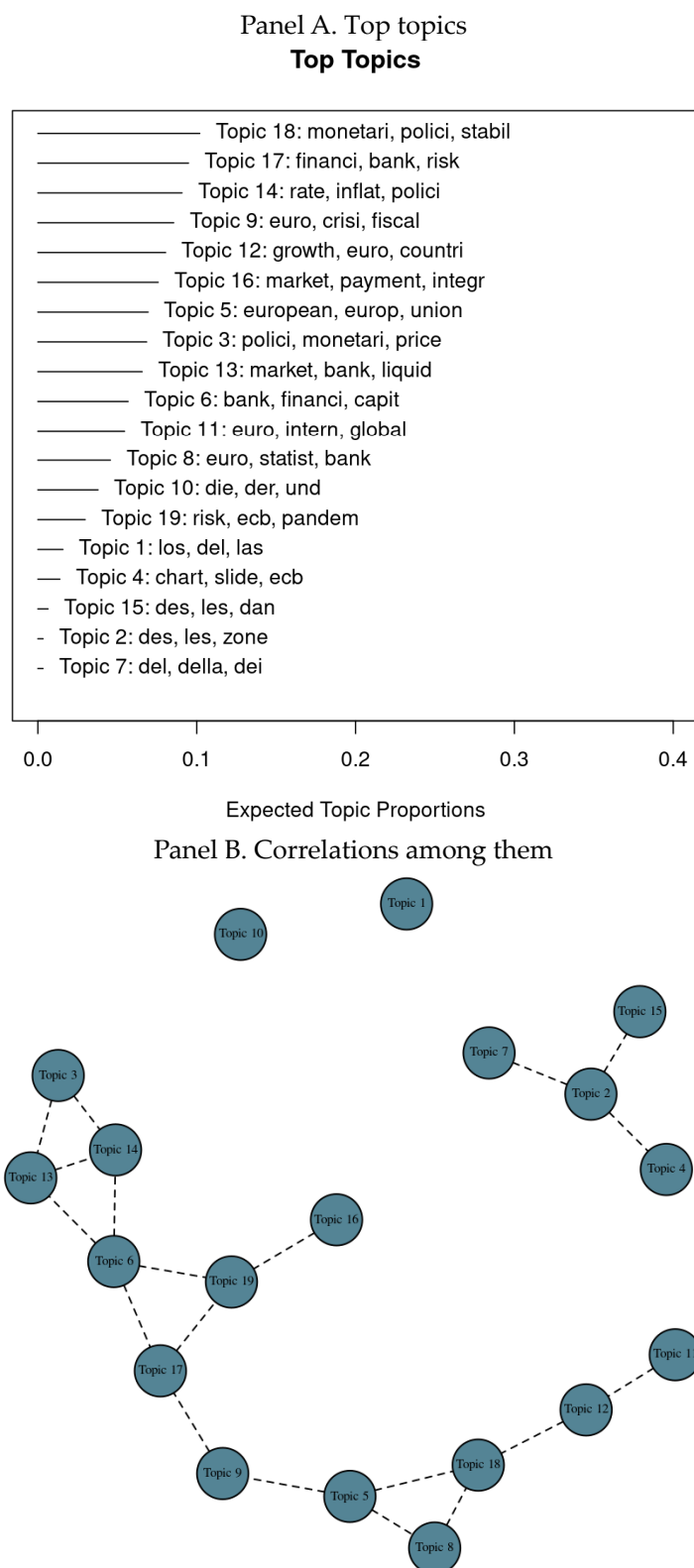**Figure 8.** Word clouds of top topics (1997–2021).

Figure 9 provides histograms of the top topics discussed in the speeches and interviews during the period of 1997–2021, after removing the frequently used definite and indefinite articles. The histograms show the frequency of each of the top topics, providing an overview of the most prevalent themes discussed by high-profile European central bankers. The most frequently discussed topics during this period were related to monetary policy and stability, financial risks and banks, inflation rates and policies, the euro crisis and fiscal issues, and growth and countries in the eurozone. The histograms highlight the significance of these topics in the communication of the European Central Bank officials over the years.
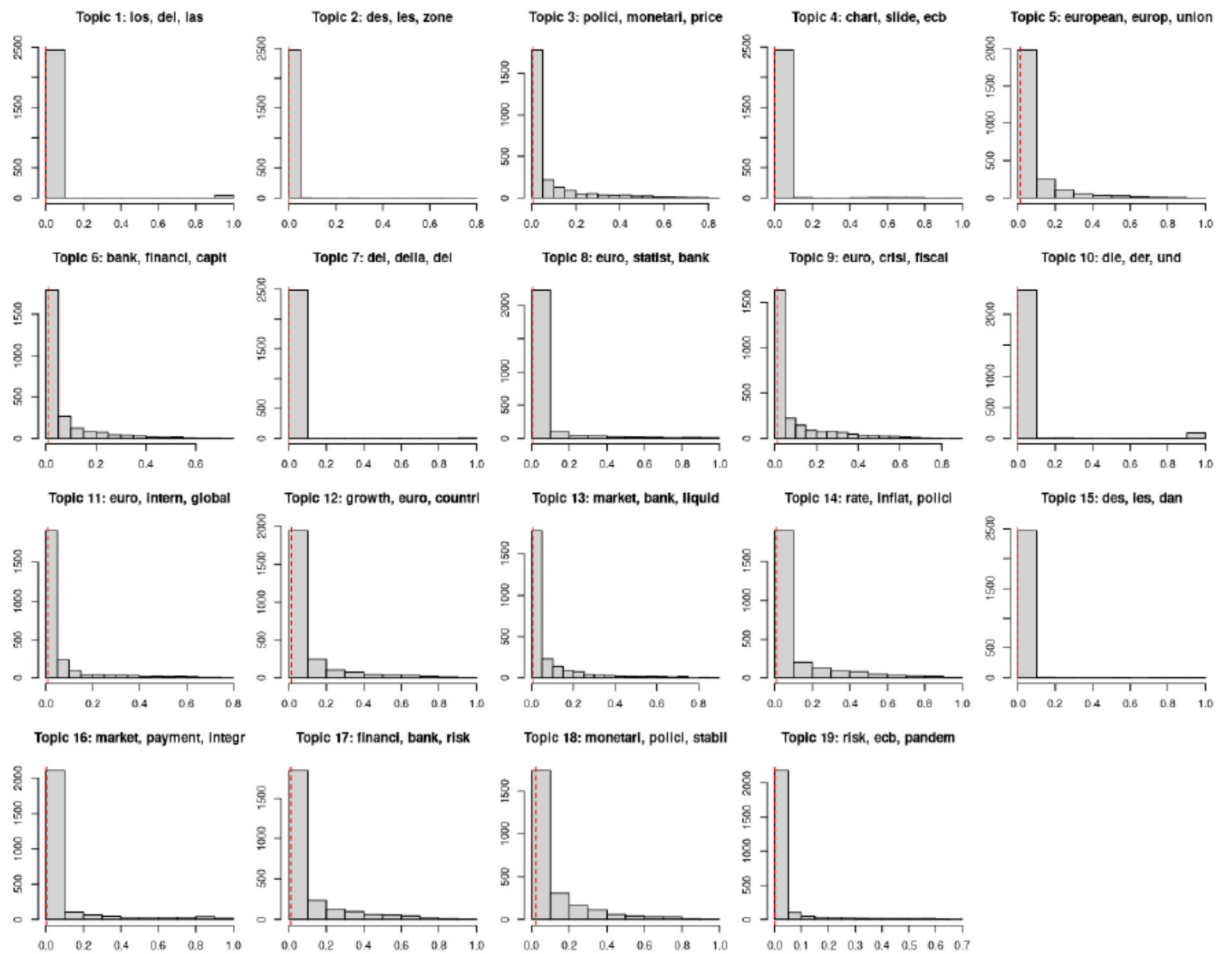
Figure 9. Histogram of top topics (1997–2021).

Figure 10 presents the results of the topical prevalence contrast analysis, which investigates the topical content of speeches in relation to the degree of central bank independence (low, medium, or high) in monetary policy. The analysis compares the expected proportion of each topic across the three categories of central bank independence and highlights the topics that are more prevalent in one category compared to the others. These results provide insights into how the level of central bank independence affects the topics discussed in ECB officials' speeches and interviews.

The semantic analysis of the speeches and interviews of high-profile European central bankers reveals interesting insights into the relationship between the level of central bank independence and the topics discussed. The analysis shows that the most frequently used topic [T-18], monetary, policy, stability, is used more often by highly independent central banks (C3) than by less independent central banks (C2 and C1). This suggests that more independent central banks are more likely to use stability topics in their official rhetoric to communicate their commitments and manage private sector expectations.

In contrast, the top topic [T-17], financial, bank, risk, and the topic [T-14] rate, inflation, policy are more frequently discussed by officials of less independent central banks. This suggests that top officials of less independent central banks tend to focus more on monetary tools and policies, while high-profile bankers of more independent central banks tend to discuss monetary objectives more.

However, it should be noted that the extent to which central bank independence affects the degree of fulfillment of main objectives is a complex issue that requires further investigation. The results presented in this study provide a starting point for exploring the

relationship between central bank independence and the topics discussed by high-profile European central bankers.
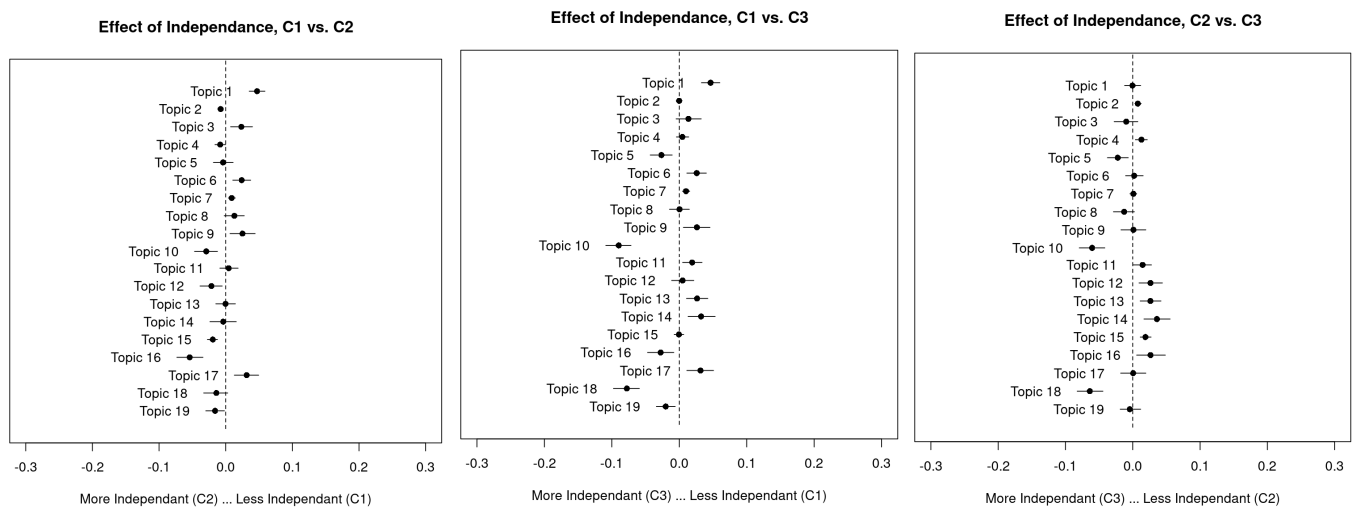

Figure 10. Topical prevalence contrast (1997–2021).

5.2. The Global Financial Crisis (2008) and the Pandemic Crisis (2020–2021)

In the context of the pandemic crisis, there was a shift towards topics related to monetary policy, stability, and economic recovery, with increased attention paid to growth, eurozone countries, and fiscal policies (Figure 7). The pandemic crisis has affected economies globally, and the speeches and interviews of European central bankers reflect the focus on addressing the economic challenges and providing stability during uncertain times. The shift towards discussing growth and fiscal policies suggests a shift towards a more expansionary monetary policy to support economic recovery. The analysis of semantic changes during crises highlights the importance of monitoring the communication of central bankers in response to changing economic conditions and crises.

Figure 11 shows the expected proportions of the top topics during the 2008 Global Financial Crisis, a period marked by significant economic challenges and financial instability. The analysis highlights the prevalence of topics related to financial risks, market volatility, and the stability of the financial system. Specifically, topics such as [T-17] financial, bank, risk, [T-13] market, bank, liquidity, and [T-6] bank, financial, capital were more prevalent during this period, indicating the increased attention given to financial risks and the stability of the banking system. In contrast, topics related to European integration ([T-5] European, Union, Europe) and inflation ([T-16] inflation, rate, policy) were less-addressed topics in bankers' speeches and interviews. Overall, Figure 11 provides insight into the changes in the rhetoric of European central bankers during a significant economic crisis.

The importance of risks was highlighted in an interview by Mr. Jean-Claude Trichet, President of the European Central Bank, for Süddeutsche Zeitung on 29 April 2009, emphasizing that: "the quantity of risk and the price of risks were underestimated in global finance. Spreads were very low, risk premia were abnormally meagre, and volatility was very modest. This called for a market correction, and we asked market participants to prepare for that market correction" (Trichet 2009). This quote from Mr. Jean-Claude Trichet highlights the importance of risks in the global financial system during the 2008 Global Financial Crisis. Mr. Trichet notes that the quantity and price of risks were underestimated, and that there was a need for a market correction. He also mentions that market participants were asked to prepare for the correction. This quote underscores the importance of monitoring risks in the financial system and the need for regulators and policymakers to take action to mitigate them.
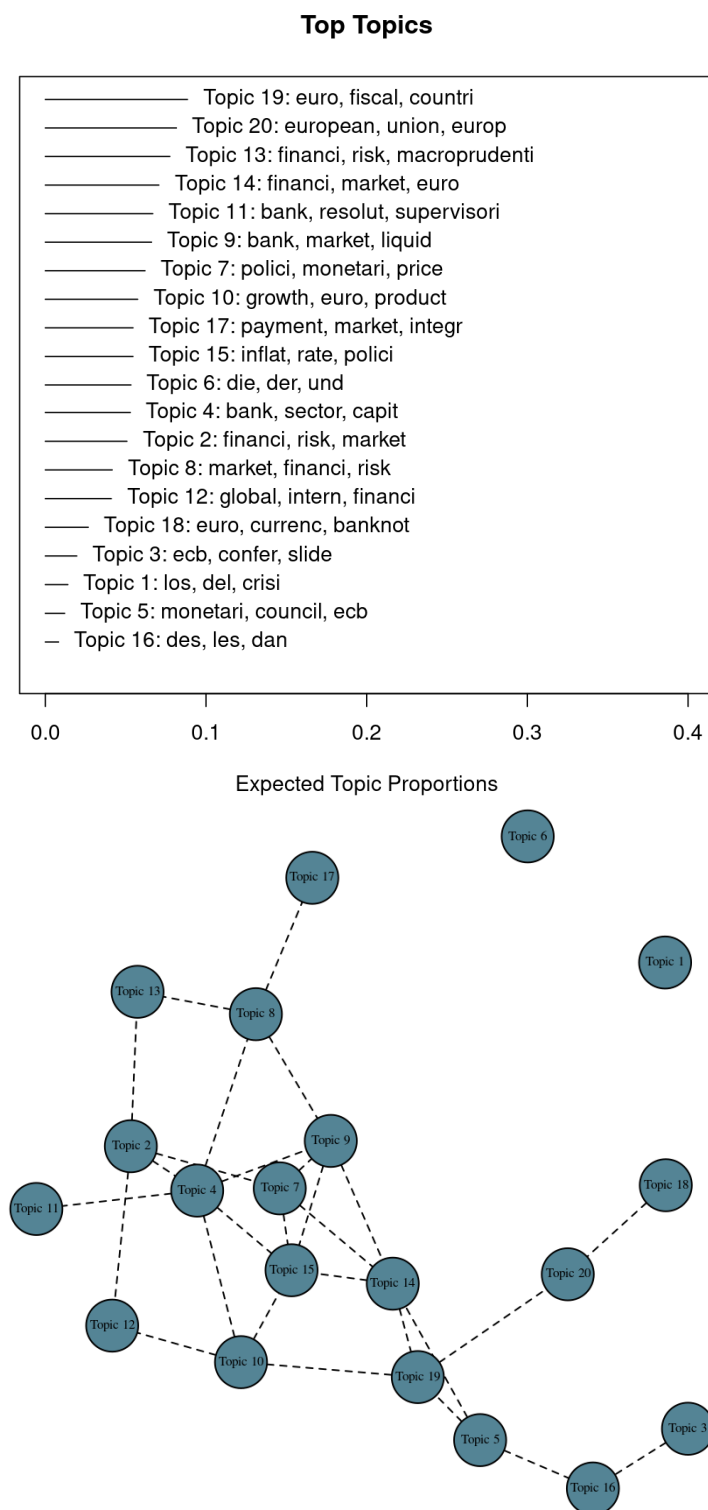
Top Topics

───────────	Topic 19: euro, fiscal, countri
───────────	Topic 20: european, union, europ
───────────	Topic 13: financi, risk, macroprudenti
───────────	Topic 14: financi, market, euro
───────────	Topic 11: bank, resolut, supervisori
───────────	Topic 9: bank, market, liquid
───────────	Topic 7: polici, monetari, price
───────────	Topic 10: growth, euro, product
───────────	Topic 17: payment, market, integr
───────────	Topic 15: inflat, rate, polici
───────────	Topic 6: die, der, und
───────────	Topic 4: bank, sector, capit
───────────	Topic 2: financi, risk, market
───────────	Topic 8: market, financi, risk
───────────	Topic 12: global, intern, financi
───────────	Topic 18: euro, currenc, banknot
───────────	Topic 3: ecb, confer, slide
───────────	Topic 1: los, del, crisi
───────────	Topic 5: monetari, council, ecb
───────────	Topic 16: des, les, dan

0.0   0.1   0.2   0.3   0.4

Expected Topic Proportions

Figure 11. Expected topic proportions during the 2008 Global Financial Crisis.

Figure 12 illustrates the topical prevalence contrast during the 2008 Global Financial Crisis, comparing the expected proportion of each topic across the three categories of central bank independence in monetary policy. The analysis highlights the topics that are more prevalent in one category compared to the others, providing insights into how the level of central bank independence affects the topics discussed in ECB officials' speeches and interviews during this crisis period. The most notable difference is the prevalence of the

[T-17] financial, bank, risk topic, which is more prevalent in the speeches and interviews of less independent central banks (C2 and C1) compared to more independent central banks (C3). This suggests that during the crisis, less independent central banks were more focused on financial risks and the stability of the banking system. In contrast, topics related to [T-18] monetary policy and stability and [T-16] inflation, rate, policy) were more prevalent in the speeches and interviews of more independent central banks (C3). Overall, Figure 12 provides insights into how the level of central bank independence can influence the topics discussed during a significant economic crisis.
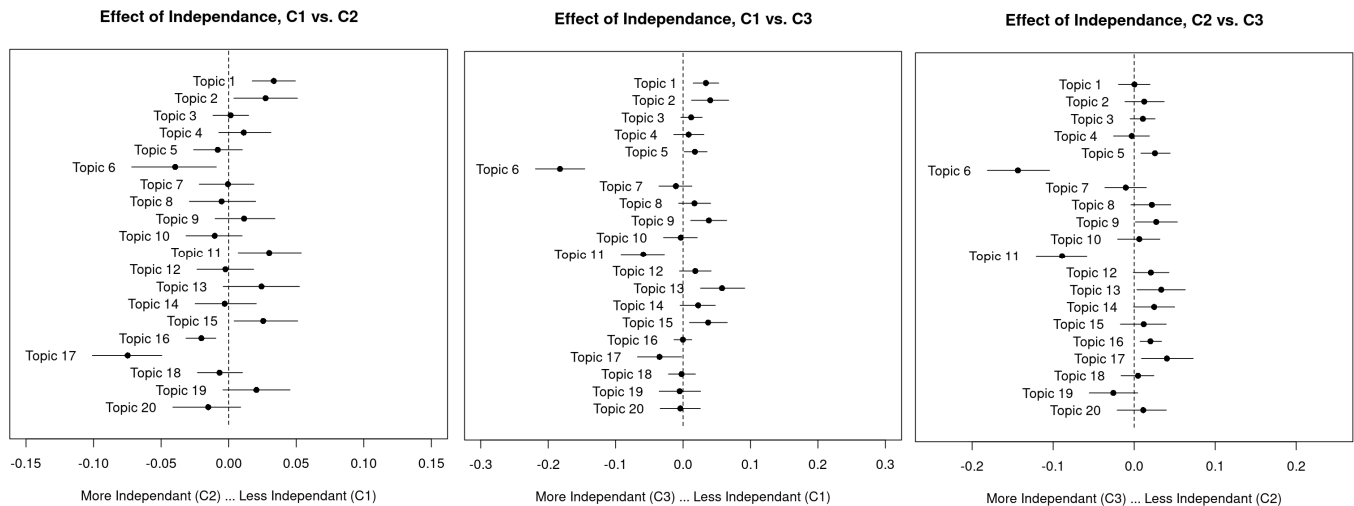
Figure 12. Topical prevalence contrast (Global Financial Crisis 2008).

We also accentuate the important role of the ECB during the pandemic crisis, as evidenced by the relatively high position of the topic monetary, Council, and ECB. The pandemic-induced recession required unconventional and unprecedented support from the monetary policy. Central banks shouldered the burden of the crisis by providing liquidity support. It is not surprising, therefore, that policy, support, and pandemic; policy, conditions, monetary; and policy, rate, monetary were the most frequently used topics in European central bankers' speeches and interviews. The major concerns were neatly addressed by Philip Lane, member of the Executive Board of the European Central Bank in August 2020: "The nature of the pandemic shock called for an extraordinary policy response. From the outset, there were three challenges for the ECB: (i) to stabilise markets; (ii) to protect credit supply; and (iii) to neutralise the pandemic-related downside risks to the inflation path. Tackling the first pair of challenges is needed in order to achieve the inflation aim, since it is problematic to run an effective monetary policy under conditions of market instability or a credit crunch" (Lane 2020).

The crisis was also perceived as an opportunity. Central bankers' speeches and interviews linked the pandemic crisis to the opportunities to tackle topics as green, market, investment as well as climate, change, risk.

Figure 13 displays the expected proportions of the top topics during the pandemic crisis period (1st February 2020–20th October 2021). The analysis shows a shift towards topics related to monetary policy, stability, and economic recovery, with increased attention given to growth, eurozone countries, and fiscal policies. The most frequently used topics during this period were related to policy support and pandemic, policy conditions and monetary, and policy rate and monetary. This indicates that European central bankers were focused on addressing the economic challenges and providing stability during the uncertain times brought about by the pandemic. The analysis of expected topic proportions during pandemic times highlights the important role of central banks in providing support during economic crises and the need for monitoring their communication during such periods.
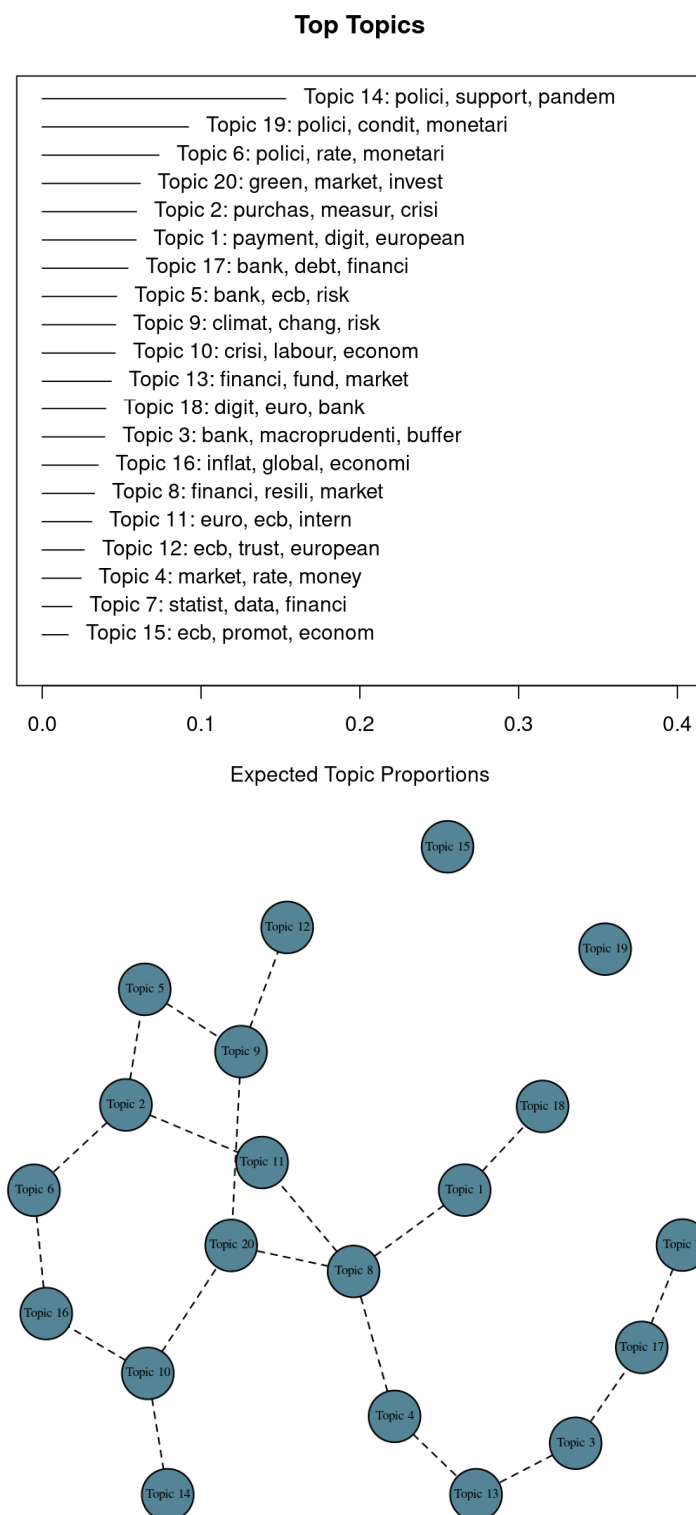
Top Topics

Figure 13. Expected topic proportions during pandemic times (2020–2021).

The most frequently used topics during this period were related to policy support and pandemic, policy conditions and monetary, and policy rate and monetary. This indicates that European central bankers were focused on addressing the economic challenges and providing stability during the uncertain times brought about by the pandemic. The analysis of expected topic proportions during pandemic times highlights the important role of

central banks in providing support during economic crises and the need for monitoring their communication during such periods.

Figure 14 displays the results of the topical prevalence contrast analysis during the pandemic crisis, comparing the expected proportion of each topic across the three categories of central bank independence in monetary policy. The analysis highlights the topics that are more prevalent in one category compared to the others, providing insights into how the level of central bank independence affects the topics discussed in ECB officials' speeches and interviews during the pandemic crisis. The figure shows comparisons between C1 and C2, C1 and C3, and C2 and C3.
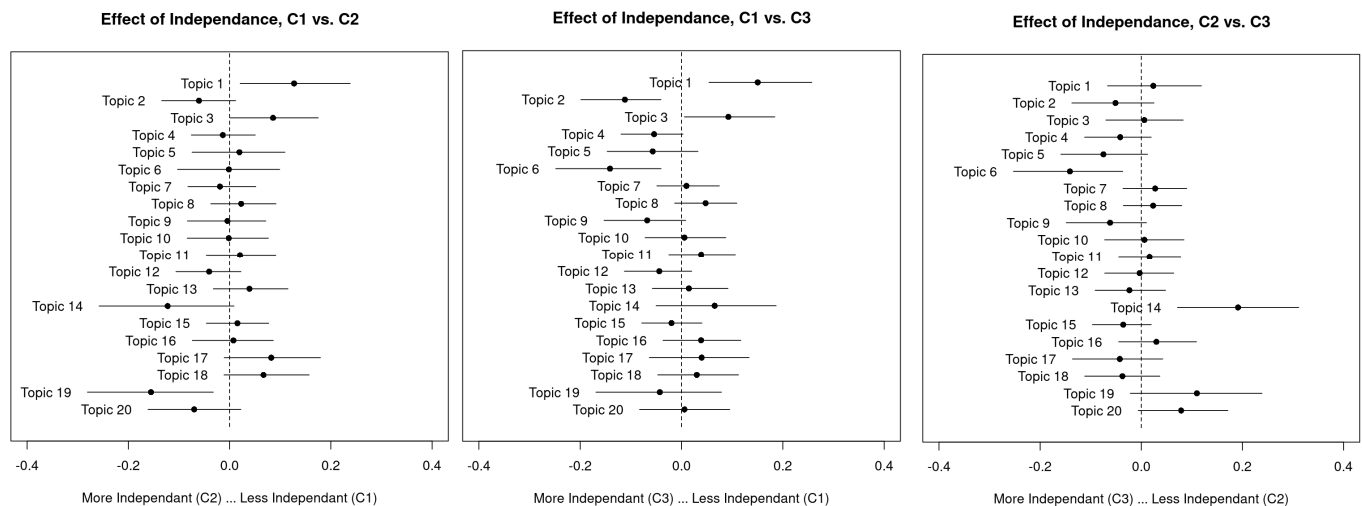
Figure 14. Topic prevalence contrast during the pandemic crisis.

The results reveal that topics related to monetary policy, such as [T-18] monetary, policy, stability, and [T-14] rate, inflation, policy, were more prevalent in the speeches and interviews of officials from more independent central banks (C3) compared to less independent central banks (C1 and C2). In contrast, topics related to financial risks, such as [T-17] financial, bank, risk, and [T-13] market, bank, liquidity, were more prevalent in the speeches and interviews of officials from less independent central banks (C1 and C2) compared to more independent central banks (C3). These results suggest that central bank independence plays a significant role in shaping the topics discussed in ECB officials' speeches and interviews during crises, such as the pandemic crisis.

5.3. Periods of Heightened Inflation

We identify episodes or periods of heightened inflation by using the monthly inflation rates that are one standard deviation or more above the long-term average inflation rate for the 1997–2021 period. The speeches and interviews delivered during these periods are focused on [T-2] inflation, price, rate; [T-8] climate, ECB, policy; and [T-12] policy, euro, fiscal (Figure 15). To illustrate the defining spirit of those periods, we refer to the remarks by Ms. Christine Lagarde, President of ECB, in a speech delivered on 13 June 2020: "Dealing with uncertainty about the outlook is the task of stabilisation policies, and here the ECB will continue to take all necessary steps within our mandate of price stability—for all euro area citizens in all parts of the euro area" (Lagarde 2020).
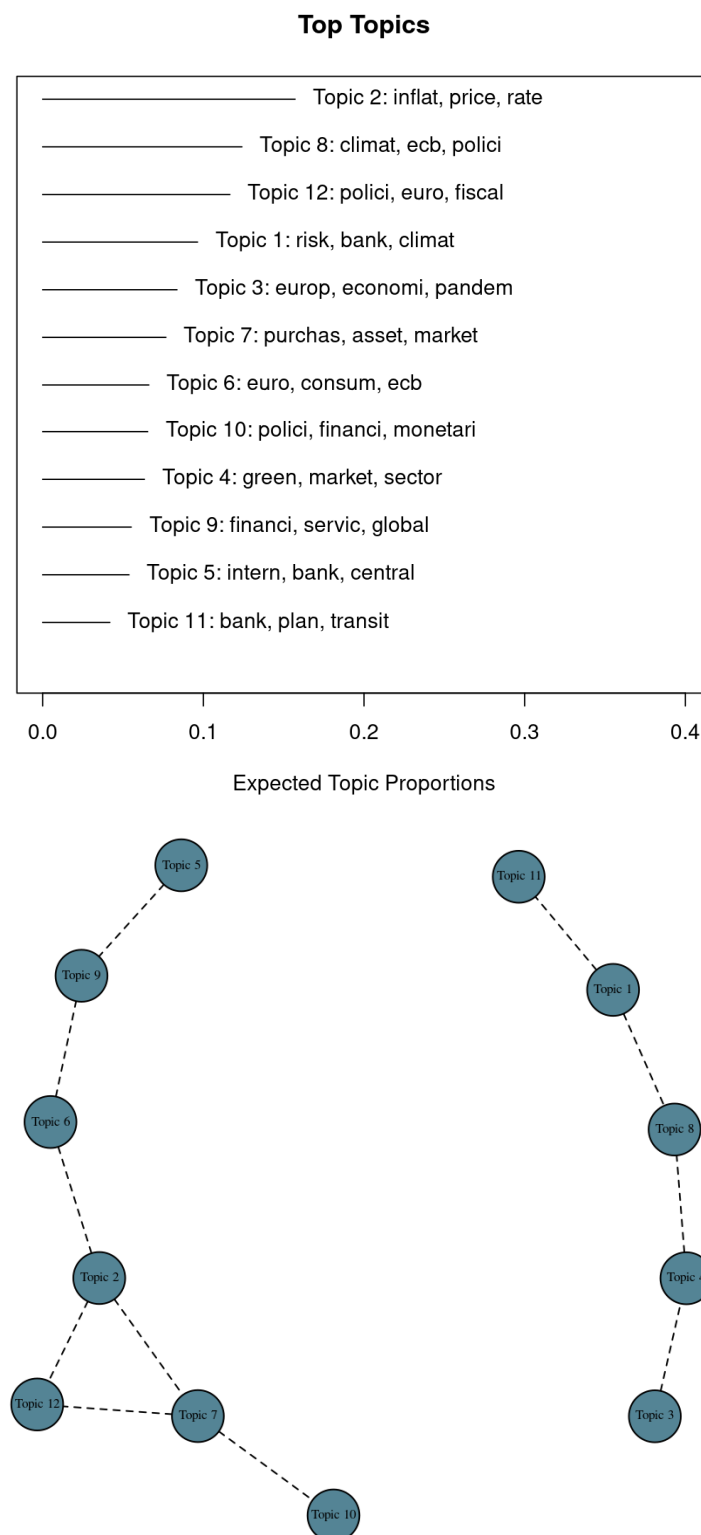
Top Topics

Figure 15. Expected topic proportions during periods of heightened inflation.

Figure 15 displays the expected topic proportions during periods of heightened inflation, identified by monthly inflation rates that are one standard deviation or more above the long-term average inflation rate for the 1997–2021 period. The analysis highlights the prevalence of topics related to inflation rates and policies, climate change and ECB policy, and policy and fiscal issues in these speeches and interviews. The focus on inflation is expected during these periods, as central banks typically prioritize price stability. However, there is

also a notable shift towards discussing the role of fiscal policy in addressing inflation and the need for monetary policy tools, such as interest rates, to manage inflationary pressures.

A notable semantic change is the refocus towards monetary policy tools ("rates") and the need for the fiscal policy to also shoulder the burden of the crisis. In terms of topic prevalence contrast, we underline that more independent central banks are more inclined (compared to less independent ones) to discuss risks and climate changes and the role of ECB, even in periods of higher headline inflation rates (Figure 15).

Figure 16 displays the topic prevalence contrast during periods of heightened inflation for the various category comparisons C1 and C2, C1 and C3, and C2 and C3. The analysis highlights the prevalence of topics related to inflation, price, and rate during these periods, indicating the increased attention given to maintaining price stability. In addition, the topics of climate, ECB, and policy and policy, euro, and fiscal were also prevalent, indicating the importance of addressing climate change and the role of fiscal policy in maintaining stability during periods of heightened inflation.
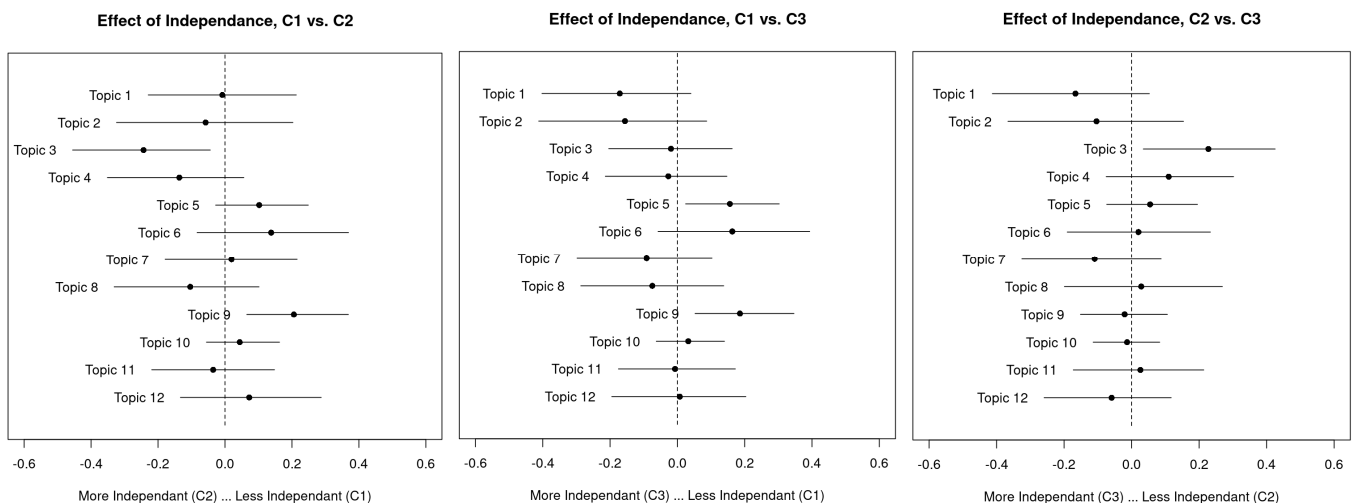
Figure 16. Topic prevalence contrast during periods of heightened inflation.

The results of the topic prevalence contrast analysis also show that more independent central banks (C3) were more likely to discuss risks, climate changes, and the role of the ECB during periods of heightened inflation compared to less independent central banks (C1 and C2). This suggests that more independent central banks may have a broader perspective on the factors that contribute to maintaining stability during times of economic turbulence. Overall, Figure 16 provides insight into the differences in the communication of European central bankers during periods of heightened inflation, based on the level of central bank independence.

5.4. Limitations of the Research

We performed a thematic analysis of the ECB communication without searching for a specific pre-specified theme. This agnostic approach does not focus on pre-defined structural topics. On the other hand, the sample period of the analysis is information-rich precisely for comprising two major financial crises (subprime and eurozone sovereign debt crises) and a global pandemic. Therefore, we acknowledge the limited focus on pre-specified themes as one of the important limitations of the research. For illustration, two specific topics that should be investigated as future research avenues are the international monetary policy coordination and the implicit government guarantees to the banking system.

International monetary policy coordination: One of the key differences between the monetary policy responses following the subprime crisis and the COVID-19 pandemic is

that the former was characterized by unilateral policy responses and the latter by multilateral actions by several central banks. This issue is thoroughly discussed in the study by Cortes et al. (2022), highlighting that several central banks expanded their balance sheets concomitantly at the onset of the pandemic. Therefore, it is important to discuss the role of monetary policy coordination—as in Cortes et al. (2022)—and analyze whether ECB speeches refer to other central banks such as the US Fed, the Bank of England, or the Bank of Japan. This is particularly relevant from the perspective of the unconventional monetary policy announcements around the world.

Implicit government guarantees to the banking sector: During financial crises, implicit government guarantees are critically relevant to ensure the stability of the banking sector. In the context of the eurozone, the ECB may function as the lender of last resort and alleviate concerns that the banking sectors of the eurozone member countries collapse. For example, Dantas et al. (2023) explore why the eurozone itself creates an extra layer of implicit guarantees granted to banks. Accordingly, one would expect terms such as "bank bailouts" to appear in ECB speeches during the eurozone sovereign debt crisis. A more nuanced analysis would uncover whether "bank bailouts" as a term was more used in the speeches to reflect difficult trade-offs in the past (ex post reflection) or as an argument to strengthen the bank capital and prevent frequency or severity of future crises (forward guidance).

6. Conclusions

In conclusion, our study demonstrates the value of using quantitative indices and text mining techniques to better understand central bank communication and transparency. By employing Structural Topic Modeling (STM) and incorporating metadata, we were able to uncover prominent topics in the speeches and interviews of high-profile European central bankers from 1997 to 2021. Our analysis included different periods of crisis and periods of heightened inflation, which allowed us to identify notable semantic changes and different emphases in the speeches and interviews of central bankers.

One of our key findings is that more independent central banks tend to focus more on monetary goals, while less independent central banks tend to discuss monetary policy tools. This observation sheds light on the potential differences in the official rhetoric of central banks with varying degrees of independence. We also found notable changes in the topics discussed during crisis periods, such as the shift towards discussing the role of institutions during the pandemic crisis and the refocus on monetary policy tools and the need for fiscal policy to play a role during periods of heightened inflation.

Our study has important implications for policymakers, financial markets, and researchers. The insights gained from our analysis can help anticipate policy decisions, assess the effectiveness of communication strategies, and explore the implications of increased transparency in modern monetary policy. For example, policymakers could use this information to adjust their communication strategies, while financial markets could better anticipate policy decisions. Researchers could further explore the impact of central bank communication on financial markets and focus on forward guidance.

In conclusion, our study highlights the value of text mining techniques in enhancing our understanding of central bank communication and transparency. Future research could extend this analysis to other central banks or incorporate additional data sources to gain even deeper insights into central bankers' communication and its impact on financial markets and the broader economy.

References

Apel, Mikael, Marianna B. Grimaldi, and Isaiah Hull. 2022. How Much Information Do Monetary Policy Committees Disclose? Evidence from the FOMC's Minutes and Transcripts. *Journal of Money, Credit and Banking* 54: 1459–90. [CrossRef]

Bank for International Settlements. 2023. Central Bankers' Speeches. Available online: https://www.bis.org/cbspeeches/index.htm (accessed on 11 January 2022).

Bennani, Hamza. 2019. Does People's Bank of China Communication Matter? Evidence from Stock Market Reaction. *Emerging Markets Review* 40: 100617. [CrossRef]

Bernanke, Ben. 2004. The Great Moderation. In *Taylor Rule and the Transformation of Monetary Policy*. Edited by Evan F. Koenig. Washington, DC: Hoover Institute Press.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2019. *Monetary Policy Communications and Their Effects on Household Inflation Expectations*. NBER Working Paper, No. 25482. Cambridge, MA: National Bureau of Economic Research. [CrossRef]

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2022. Monetary Policy Communications and Their Effects on Household Inflation Expectations. *Journal of Political Economy* 130: 1537–84. [CrossRef]

Cortes, Gustavo S., George P. Gao, Felipe B. Silva, and Zhaogang Song. 2022. Unconventional Monetary Policy and Disaster Risk: Evidence from the Subprime and COVID-19 Crises. *Journal of International Money and Finance* 122: 102543. [CrossRef] [PubMed]

Cortes, Gustavo S., Mani Sethuraman, and Felipe B. G. Silva. 2023. Monetary Policy and Corporate Communication: Evidence from Conference Calls. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3798794 (accessed on 21 January 2023).

Dantas, Manuela, Kenneth J. Merkley, and Felipe B. G. Silva. 2023. Government Guarantees and Banks' Income Smoothing. *Journal of Financial Services Research* 63: 123–73. [CrossRef]

Dedola, Luca, Georgios Georgiadis, Johannes Gräb, and Arnaud Mehl. 2020. Does a Big Bazooka Matter? Quantitative Easing Policies and Exchange Rates. *Journal of Monetary Economics* 117: 489–506. [CrossRef]

Dincer, Nergiz, and Barry Eichengreen. 2014. Central Bank Transparency and Independence: Updates and New Measures. *International Journal of Central Banking* 10: 189–259. [CrossRef]

Dincer, Nergiz, Barry Eichengreen, and Petra M. Geraats. 2019. Transparency of Monetary Policy in the Postcrisis World. In *The Oxford Handbook of the Economics of Central Banking*. Edited by David Mayes, Pierre Siklos and Jan-Egbert Sturm. Oxford: Oxford University Press, chp. 10.

Dincer, Nergiz, Barry Eichengreen, and Petra M. Geraats. 2022. Trends in Monetary Policy Transparency: Further Updates. *International Journal of Central Banking* 18: 331–48.

Eijffinger, Sylvester, and Petra M. Geraats. 2006. How transparent are central banks? *European Journal of Political Economy* 22: 1–21. [CrossRef]

Garriga, Ana Carolina. 2016. Central Bank Independence in the World: A New Dataset. *International Interactions* 42: 849–68. [CrossRef]

Geraats, Petra M. 2002. Central Bank Transparency. *The Economic Journal* 112: F532–F565. [CrossRef]

Geraats, Petra M. 2020. The ECB's Monetary Policy Strategy: Lessons from the Financial Crisis, Debt Crisis and Coronavirus Crisis. In *Contributions to the Strategy Review of the European Central Bank: A Report from The ECB and Its Watchers XXI Conference*. IMFS Interdisciplinary Studies in Monetary and Financial Stability 1/2020. Edited by Volker Wieland. Frankfurt: Institute for Monetary and Financial Stability, pp. 79–91.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera. 2023. The Voice of Monetary Policy. *American Economic Review* 113: 548–84. [CrossRef]

Grilli, Vittorio, Donato Masciandaro, and Guido Tabellini. 1991. Political and Monetary Institutions and Public Financial Policies in the Industrial Countries. *Economic Policy* 6: 341–92. [CrossRef]

Hansen, Stephen, and Michael McMahon. 2016. Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication. *Journal of International Economics* 99: S114–S133. [CrossRef]

Hansen, Stephen, Michael Mcmahon, and Andrea Prat. 2018. Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. *Quarterly Journal of Economics* 133: 801–70. [CrossRef]

Hansen, Stephen, Michael McMahon, and Matthew Tong. 2019. The Long-Run Information Effect of Central Bank Communication. *Journal of Monetary Economics* 108: 185–202. [CrossRef]

Hubert, Paul, and Fabien Labondance. 2017. *Central Bank Sentiment and Policy Expectations*. Bank of England Working Paper No. 648. London: Bank of England.

Issing, Otmar. 2005. Communication, Transparency, Accountability: Monetary Policy in the Twenty-First Century. *Federal Reserve Bank of St. Louis Review* 87: 65–83. [CrossRef]

Kahveci, Eyup, and Aysun Odabaş. 2016. Central Banks' Communication Strategy and Content Analysis of Monetary Policy Statements: The Case of Fed, ECB and CBRT. *Procedia—Social and Behavioral Sciences* 235: 618–29. [CrossRef]

Küsters, Anselm. 2022. Applying Lessons from the Past? Exploring Historical Analogies in ECB Speeches through Text Mining, 1997–2019. *International Journal of Central Banking* 18: 277–329. [CrossRef]

Lagarde, Christine. 2020. The Path Out of Uncertainty. Remarks by Ms Christine Lagarde, President of the European Central Bank, at the Inaugural Session of the Italian National Consultation, 13 June 2020 (Videoconference). Available online: https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200613~890270bad1.en.html (accessed on 21 January 2023).

Lane, Philip. 2020. The Pandemic Emergency—The Three Challenges for the ECB. Speech at "Navigating the Decade Ahead: Implications for Monetary Policy", A Symposium, Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, 27 August 2020. Federal Reserve Bank, Kansas City (USA). Available online: https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200827~1957819fff.en.html (accessed on 21 January 2023).

Maqsood, Haider, Irfan Mehmood, Muazzam Maqsood, Muhammad Yasir, Sitara Afzal, Farhan Aadil, Mahmoud M. Selim, and Khan Muhammad. 2020. A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management* 50: 432–51. [CrossRef]

Masawi, Becksndale, Sukanto Bhattacharya, and Terry Boulter. 2014. The Power of Words: A Content Analytical Approach Examining whether Central Bank Speeches Become Financial News. *Journal of Information Science* 40: 198–210. [CrossRef]

Mathur, Aakriti, and Rajeswari Sengupta. 2019. *Analysing Monetary Policy Statements of the Reserve Bank of India*. IHEID Working Papers 08-2019. Geneva: Economics Section, The Graduate Institute of International Studies.

Morris, Stephen, and Hyun S. Shin. 2002. Social Value of Public Information. *The American Economic Review* 92: 1521–34. [CrossRef]

Petropoulos, Anastasios, and Vasilis Siakoulis. 2021. Can Central Bank Speeches Predict Financial Market Turbulence? Evidence from an Adaptive NLP Sentiment Index Analysis Using XGBoost Machine Learning Technique. *Central Bank Review* 21: 141–53. [CrossRef]

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2019. stm: An R Package for Structural Topic Models. *Journal of Statistical Software* 91: 1–40. [CrossRef]

Sapphasak, Chatchawan. 2019. *Words Matter: Effects of Semantic Similarity of Monetary Policy Committee's Decision on Financial Market Volatility*. PIER Discussion Papers 121. Bangkok: Puey Ungphakorn Institute for Economic Research.

Shapiro, Adam H., and Daniel Wilson. 2021. *Taking the Fed at Its Word: A New Approach to Estimating Central Bank Objectives Using Text Analysis*. Federal Reserve Bank of San Francisco Working Paper 2019-02. San Francisco: Federal Reserve Bank. [CrossRef]

Su, Shiwei, Ahmad H. Ahmad, and Justine Wood. 2019. How effective is central bank communication in emerging economies? An empirical analysis of the Chinese money markets responses to the people's bank of China's policy communications. *Review of Quantitative Finance and Accounting* 54: 1195–219. [CrossRef]

Summers, Lawrence. 2020. *Accepting the Reality of Secular Stagnation. Point of View*. Washington, DC: Finance & Development, The International Monetary Fund, pp. 17–19.

Taddy, Matt. 2015. Distributed multinomial regression. *Annals of Applied Statistics* 9: 1394–414. [CrossRef]

Trichet, Jean-Claude. 2009. Interview with Mr Jean-Claude Trichet, President of the European Central Bank, and Süddeutsche Zeitung, Conducted by Ms Helga Einecke and Mr Martin Hesse, Published on 29 April 2009. Süddeutsche Zeitung, Munich. Available online: https://www.bis.org/review/r090429a.pdf (accessed on 21 January 2023).

Wallach, Hanna, Iain Murray, Ruslan Salakhutdinov, and David Mimno. 2009. Evaluation Methods for Topic Models. Paper presented at the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, June 14–18; New York: Association for Computing Machinery, pp. 1105–12. [CrossRef]

Warin, Thierry, and William Sanger. 2020. How Data Science can (also) help central bankers: An NLP study of the European Central Bank presidents' speeches. *Global Economy Journal* 20: 2050009. [CrossRef]

Woodford, Michael. 2005. *Central Bank Communication and Policy Effectiveness*. Proceedings—Economic Policy Symposium—Jackson Hole. Kansas City: Federal Reserve Bank of Kansas City, pp. 399–474.