

The diffusion of topics in central bank speeches*

Dave Campbell^a, James Chapman^b, Jon Frost^c, Carola Müller^d, and Andrea Payne^a

^a*Carleton University*

^b*Bank of Canada*

^c*Bank for International Settlements (BIS)*

^d*Banco de España*

— Preliminary and incomplete draft —

Abstract

We document the spread of economic narratives by analysing topics in central bank speeches. We use natural language processing to identify topics and how they evolve over time in the speeches of senior officials from G20 central banks and the Bank for International Settlements (BIS). We then test for the presence of thought leaders among central banks by using lasso regression for variable selection. We find that central banks commonly pay attention their peers. We can identify thought leaders for the topic of macroprudential policy but not for monetary policy. At most, some regional dialogue seems to matter when talking about monetary policy or inflation.

JEL Classification: E58, G28.

Keywords: Central bank communication, central bank speeches, topic modelling.

*The views expressed here are those of the authors and do not necessarily reflect those of the BIS, Bank of Canada, Bank of Spain or the Eurosystem. We thank Fernando Perez-Cruz and Taejin Park for helpful comments. All errors are our own.

1 Introduction

Central banks use communication to manage expectations. This, in turn, shapes economic outcomes (Blinder, Ehrmann, Fratzscher, De Haan, and Jansen, 2008). This is most visible when central banks use forward guidance as a policy instrument. By communicating their assessment of the economy, policy intentions and outlook, central banks provide their interpretation of economic and financial developments, including on novel topics. These narratives can spread through the public or academic sphere (Shiller, 2017). In a globalised economy, with a highly interconnected financial system, monetary policy in jurisdictions with key financial markets spills over to other economies (Bruno and Shin, 2015; Aizenman, Chinn, and Ito, 2016; Miranda-Agrippino and Rey, 2020). This makes it imperative for central banks to listen to one another, and for researchers to understand how the diffusion of topics across central bank works.

We conjecture that narratives around specific new topics spread through central banks, that certain central banks are particularly influential and that narratives are reflected in the public speeches that central bank governors and board members give.

In this paper, we document the spread of topics across central banks through central bankers' speeches. We use natural language processing (NLP) to identify topics and how they evolve over time in the speeches of senior officials from the Group of 20 (G20) central banks. We then use lasso regression to further test for the presence of *thought leaders*. We conceive of thought leaders as those central banks that start and promote certain narratives. Technically, we identify thought leaders as central banks whose mentioning of topics in speeches consistently increases the likelihood of other central banks mentioning the same topic thereafter. Of course, given their narrow mandates and focus on their domestic economy, it is not clear *a priori* that central banks should influence one another. But given regular dialogue and cooperation with one another, there is the potential for thought leaders to share their perspectives with their peers and thus for certain topics to diffuse internationally. By studying how central banks influence one another, we contribute to the understanding of central bank communication, bringing in an important cross-border dimension.

In our analysis, we proceed in two steps. First, we use non-negative matrix factorisation (NMF) to identify topics in speeches. We use 11,360 speeches by governors or board members of G20 central banks, including the European Central Bank (ECB), and the Bank for International Settlements (BIS), an international organisation that serves the central bank community. We limit the NMF to 40 topics. The NMF assigns a probability to each topic within each speech: the topic proportion. Topic proportions give the likelihood that the speech was about that topic relative to all other topics. From this, we construct average topic proportions (ATP) per month, topic and institution. For example, the ATP of the BIS in November 2023 for the topic “digital.currency_” is 0.093, putting digital currency among the four most mentioned topics by the BIS in that month.¹

In the next step, we build on this panel of 21 institutions and 40 topics per month to estimate thought leaders using lasso regressions for variable selection. We match macro-financial variables to the institutions based on their jurisdiction, with exception of the ECB, which represents 21 economies (of which three are G20 members) and the BIS (which has 63 central bank shareholders, including all G20 members). We conjecture that an institution’s topic choice is driven by the current macro-financial environment represented by the covariates and potentially topics chosen by thought leaders in the near past represented by lagged ATPs of other institutions. In the corresponding regression, we let the lasso choose which variables to include and thus whose lagged topic choices matter for current topic choices of their peers. The approach allows us to be agnostic about candidates for being thought leaders.

We find that central banks persistently but with varying emphasis speak about their tasks; namely monetary policy; the payment system; reserve management; and both micro and macroprudential supervision. The topic model renders reasonable patterns. For example, during and after the Great Financial Crisis (GFC) institutions increasingly talked about crises and macroprudential policy. The topic model also captures the main recurring topics

¹In November 2023, the following speeches are examples of this topic’s preponderance in the month: Agustín Carstens (BIS) spoke on “Technological innovation: small steps and giant leaps”, “Securing the future monetary system: cyber security for central bank digital currencies” and “The future monetary system: from vision to reality”. Cecilia Skingsley (BIS) spoke on “Shaping the future financial system in the public interest”, and Alexandre Tombini (BIS) spoke on “The future of money: a possible role for central bank digital currencies and their implications”.

that represent economic variables that are inputs to central bank policy making. Among those, some are related to the real economy, like growth, inflation, productivity and trade; and others that are related to financial markets, such as credit, banking, liquidity, and the (international role of the) US dollar. Further, central banks' interest in new developments can be seen with the rising importance of topics like Covid (for the Covid-19 pandemic in 2020), digital currencies or climate change.

In addition, we find that central banks frequently pay attention to their peers. A notable exception is the topic of monetary policy. In accordance with their core mandate, central banks seem to focus on domestic developments in this area without following any international thought leader, at least with regard to the ATP frequency of discussions of monetary policy as a whole. At most, some regional dialogue seems to matter when talking about monetary policy or inflation. In contrast, macroprudential policies are discussed at an international level and we can identify the ECB as a thought leader in this area. Further, European central banks also seem to be leading in the topic of payments. Despite the global importance of the US economy and financial system, the US Federal Reserve does not consistently emerge as a thought leader on these topics.

We contribute to the novel and broad field of narrative economics using textual methods. [Shiller \(2017\)](#) gives an overview on how humans use narratives to make sense of the world and how it applies to economics. Since the subject is an elusive concept, approaches in the literature to capture narratives are diverse. Several papers use topic models to identify narratives. However, [Roos and Reccius \(2024\)](#) rightly point out that assigning a topic to a text might not suffice to identify a narrative. Similar to our work, [Larsen and Thorsrud \(2018\)](#) study the evolution and spread across borders of narratives related to business cycles. Their analysis is based on news articles and has no connection to central banks. Thus, our work contributes to the literature by highlighting a new way to study the spread of narratives in central banking.

A body of literature has emerged that applies topic analysis to central bank communication through monetary policy statements and speeches. Closest to our approach, [Feldkircher and Teliha \(2024\)](#) analyse how climate change has grown as a topic in central bank speeches.

They highlight factors, such as membership in the Network for Greening the Financial System (NGFS) and the scope of central bank mandates, that impacted the spread of the topic. Recognising that climate change is an excellent case study for our approach, we provide more generalisable findings by studying a broad set of topics over a long horizon. [Hansson \(2021\)](#), [Warin and Stojkov \(2023\)](#) and [Park, Perez-Cruz, and Shin \(2025\)](#) also identify central bank speech topics over time. In line with our results, [Park et al. \(2025\)](#) find that regional clusters as well as topic clusters emerge in central bank topics. [Warin and Stojkov \(2023\)](#) show that central bank independence and crises are factors driving central banks’ topic choice. [Hansson \(2021\)](#) finds that the economic conjuncture cannot fully explain topics’ persistence. This further hints at the importance of narratives. Two other papers are close to our approach: [Armeliuss, Bertsch, Hull, and Zhang \(2020\)](#) and [Bertsch, Hull, Lumsdaine, and Zhang \(2024\)](#) study whether the sentiment or topic of speeches have spillover effects across countries, singling out the Federal Reserve as a main source of spillovers. We add to these studies by measuring the importance of intra-group communication and international spillovers. Instead of assuming a leading role for certain banks, our study will systematically identify which central banks are leading the pack and on which topics.

Other papers that analyse the content and features of central bank speeches rely on similar methods, but their focus is not on the international spread of topics across central banks. Among these, some focus on the effect of speeches on real and financial outcomes ([Andersson, Dillén, and Sellin, 2006](#); [Born, Ehrmann, and Fratzscher, 2014](#); [Hansen, McMahon, and Prat, 2018](#); [Ahrens and McMahon, 2021](#); [Petropoulos and Siakoulis, 2021](#); [Ahrens, Erdemlioglu, McMahon, Neely, and Yang, 2025](#)). Others focus on their effectiveness in facilitating communication with the public ([Hwang, Lustenberger, and Rossi, 2023](#); [Glas and Müller, 2021](#); [Ferrara and Angino, 2022](#)).

[Siklos, Amand, and Wajda \(2018\)](#) and [Arseneau and Osada \(2023\)](#) analyse the overlap between speech topics and mandates of central banks. In a world where central banks are increasingly criticised for engaging in debates outside their core mandates (“mission creep”), having empirical evidence about how these debates are formed and evolve can provide useful background. In part, our results will also reflect the surge of certain topics,

like climate change, that are further away from core price stability mandates. Since these topics were not common at the start of our sample period but then emerged quickly, they are well suited for our analysis of their spread. However, our focus is not on the fit of topics to central bank mandates.

We also contribute to the extensive literature on central bank communication by studying how central banks influence each other. This literature established that the communication accompanying central bank policies has an impact on economic and financial outcomes by influencing agents in financial markets (Conrad and Lamla, 2010; Campbell, Evans, Fisher, Justiniano, Calomiris, and Woodford, 2012; Du, Fung, and Loveland, 2018). It also established that such communication can influence households' financial decisions (Binder, 2017).² Newer contributions to this field use text analysis to pinpoint how particular features of the communication impact markets. For example, Hansen, McMahon, and Tong (2019) show that the tone of inflation reports has an impact on long-run yields. Ehrmann and Talmi (2020) find that market volatility reacts to a central bank's drafting style and decreases when new reports resemble the previous ones. Davis and Wynne (2019) show that lengthier and more complex Federal Open Market Committee (FOMC) statements move markets more. Also for households, several papers show that better readability and reduced complexity can improve the ability of households to form inflation expectations along the desired path (Bholat, Broughton, Ter Meer, and Walczak, 2019; Kryvtsov and Petersen, 2021). However, the spillovers of communication from one of these impactful institutions to another has not been studied.

2 Data

We use the BIS central bankers' speeches database.³ Our sample period stretches from September 1996 to April 2025. We limit the sample to English language speeches associated with G20 central banks, the ECB or the BIS. This results in 11,360 unique speeches delivered by senior central bankers (governors, board members, etc).

²Blinder et al. (2008) provide a review of the early literature.

³See <https://www.bis.org/cbspeeches/index.htm>.

We extract author name and the date directly from the metadata of the speech pdf. The institution of the speaker was obtained by searching for patterns of the form “[position](of—of the) [institution name]” within the first sentences of speeches. Speeches are pre-processed by removing introductions and references. So-called N-grams (sequences of N words) were constructed from commonly occurring combinations of two or three terms filtered through relevant parts of speech tagging [Hansen et al. \(2018\)](#) to produce composite terms such as “bond_market”, “balance_sheet”, “balance_of_payments” and “international_monetary_fund”. A total of 291 bi-grams and 23 tri-grams were constructed. Stop words were removed and remaining terms were stemmed to reduce the vocabulary size. Terms used at least 5 times were retained, resulting in 20,646 unique terms. Speeches were aggregated by month to align with the time resolution of economic variables.

Of interest is how the topics of discourse evolve with respect to other countries discourse and macroeconomic data. At the jurisdiction level, economic covariates include GDP growth, inflation, the policy rate and the real exchange rate toward the US dollar. Financial variables include the value and volatility of the main stock index, credit default swap (CDS) spreads on 10-year government bonds, credit growth by banks to non-financial firms, the growth rate of the banking sector’s total assets, a financial conditions index and interbank overnight interest rates (if available). Beyond this, we have the growth rate of central bank’s assets, the countercyclical capital buffer rate, bank equity ratio to proxy for Basel Standards implementation, the number of banks and the money base.

3 Modelling speech topics

3.1 Methodology: modelling topic proportions

A probabilistic topic model (PTM) describes the generative process for word occurrences in documents, generally collected as a *bag of words* into counts of terms within documents. Documents are modelled as a composition of topics, each occurring with some probability that could be based on externalities. Words within the documents arise as random draws from the terms-within-topic distributions. The most well known PTM is latent Dirichlet

allocation (LDA) (Blei, 2012); a related set of PTMs are neural topic model (NTM) variants (Wu, Nguyen, and Luu, 2024). LDA and NTMs are well suited for producing interpretable descriptions of the contents of a corpus of documents and information retrieval; but they are well known to be dependent on the choice of random seed initialisation. As an example, table 5 shows the top 5 terms across 5 topics in LDA run with different random seeds for the central bank dataset. Terms are subject to label switching (topic 2 in seed 100 becomes topic 3 in seed 101, topic 2 in seed 110), topic rotation (topic 3 in seed 100 evolves and approximates topics 5 in seed 101 and topic 3 in seed 110) and topic materialisation and dissolution (topics 5 in seed 100 and topic 5 in seed 110 do not have analogues between seeds). Decompositions of a corpus into topic models using deep learning suffers a similar instability problem with additional interpretability challenges when documents are made up of multiple topics (Wu et al., 2024). The lack of deterministic topic models from the LDA and NTMs make them unsuitable for downstream inference such as modelling covariate effects on prevalence of topics of discourse. For estimating effects and testing hypotheses, statistics need to be a deterministic function of the data. An alternative approach is based on matrix decomposition of the term-document matrix into a term within topic matrix and a topic within document matrix. Using separable non-negative matrix factorisation (NMF) produces a unique solution whose stability makes it well suited for assessing the impact of covariates on the topics of discourse while retaining the PTM interpretation (Phelan and Campbell, 2026).

In NMF, the corpus of documents is converted into a term-document-matrix, X , where entry x_{ij} is the count of the i^{th} term within document j . Factorisation $X \approx \Phi\Theta$ decomposes the term document matrix into non-negative valued matrices, where Φ defines the importance of terms-within-topics, and Θ provides the decomposition of topics-within-document. Using a separable NMF structure allows analysts to use a mixture of purely data-driven and analyst-informed topics. The analyst-informed topics are constructed by selecting seed (anchor) words that are likely to be useful for downstream tasks such as testing hypotheses and/or model building (Phelan and Campbell, 2026). The anchors are used to construct a unique set of orthogonal topics constructed based on the level of association between anchors

and other terms. Using NMF, altering the number of topics appends or removes columns (rows) from Φ (Θ) without altering the remaining anchor terms. This stability enables statically valid inference in assessing covariate effects on prevalence of topics of discourse.

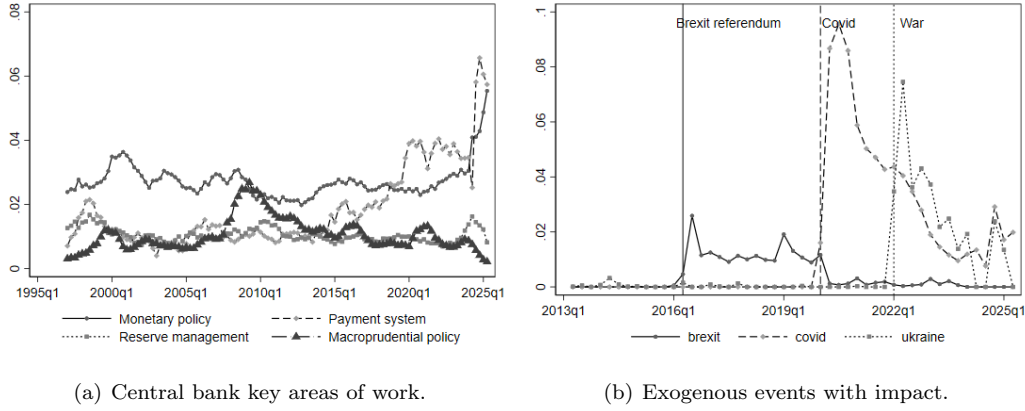
Matching the PTM interpretation, entries in the normalised Θ matrix correspond to topic within document probabilities bounded to the interval $[0, 1]$. Selected topics of analytical interest can then be modelled directly using Beta regression (Phelan and Campbell, 2026). Alternatively to perform simultaneous model selection and inference, transformed topic probabilities can be modelled using lasso for model selection and post-selection inference for confidence intervals (Tibshirani, Taylor, Lockhart, and Tibshirani, 2016).

3.2 Descriptive insights from topic proportions

We initially consider 11,360 speeches from G20 central banks alongside the the European Central Bank (ECB) and Bank for International Settlements (BIS). We run an NMF topic model with 40 topics including user defined anchors for “Ukraine”, “Brexit”, and “Covid”. These user defined anchors force these topics to appear, replacing less interpretable, more generic topics based on anchors “year”, “target”, and “develop”. Previous work has shown robustness with respect to the number of topics when considering covariate effects on topic proportions of interest Phelan and Campbell (2026). Figure 12 in the appendix provides word clouds for four key topics of interest interpretable as “monetary policy”, “macroprudential policy”, “payment system” and “reserve management”. Note that none of these were impacted by the inclusion of the custom anchors because topics in NMF are nearly orthogonal to each other. Our model identifies reasonable topics as well as topic patterns over time. For example, as can be expected, certain topics such as “inflation” or “bank” are continuously relevant for central banks. Likewise, central bank officials continuously talk about monetary policy, macroprudential policy, payment systems and reserve management as key areas of their work (see figure 1 panel a). Further, monetary policy is persistently more talked about than other central bank functions. Only during the GFC do macroprudential topics gain comparable presence in central bank speeches. Towards the end of the sample, the payment system is receiving increasing attention. From 2020 onwards this is

even more than monetary policy. Other topics, such as “Ukraine”, “Covid” or “Brexit” rise as they gain importance in international events and debates (see figure 1 panel b). In section 4.2.1, the topics are reconstructed using the 4,522 Group of 7 (G7) central bank speeches, and the corresponding word clouds are in figure 13.

Figure 1: Central bank speeches’ topic relevance over time.



Note: The lines show document topic proportions by topic. The DTP were averaged over all G-20 banks per quarter.

Furthermore, central banks tend to talk about about similar topics at the same time. The average correlation coefficient across topics and time is well above 0.5. As could be expected, central banks in more integrated regions such as Europe show stronger correlation in their topic choices than more diverse regions such as Asia-Pacific (see table 1).

Also across institutions, topics that are intrinsically connected, such as monetary policy and inflation, co-move over time (see figure 2). This is evident, for instance, during the global inflation surge in 2022-25.

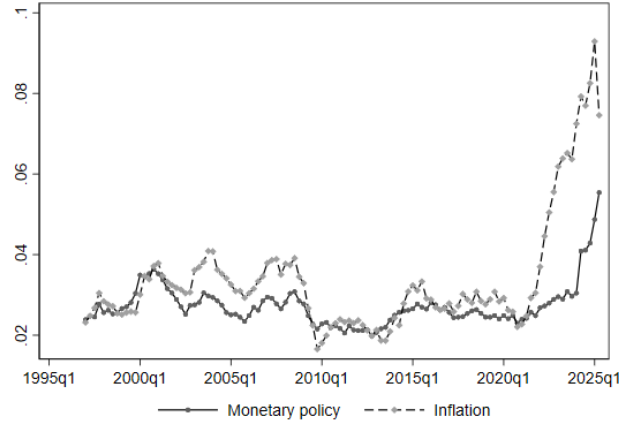
Finally, central banks started to talk about new topics during the past decade. The NMF model dynamically selects new topics over the whole sample period. In figure 3 we show the evolution of the topics “digital currency” (panel a) and “climate” (panel b) for a few institutions. Officials of some institutions such as the Bank of England mention both of these topics early on in the sample period, and others follow. Key events may be driving central banks to open a discourse on these topics. For example, following the Paris Agreement

Table 1: Correlation matrix of ATPs in different regions.

<i>Americas</i>							
	BIS	US	CA	BR	MX		
BIS	1.00						
US	0.71	1.00					
CA	0.43	0.52	1.00				
BR	0.67	0.60	0.37	1.00			
MX	0.71	0.81	0.52	0.55	1.00		
<hr/>							
<i>Europe</i>							
	BIS	EA	DE	FR	GB	IT	
BIS	1.00						
EA	0.55	1.00					
DE	0.67	0.74	1.00				
FR	0.52	0.70	0.71	1.00			
GB	0.66	0.62	0.75	0.59	1.00		
IT	0.52	0.58	0.67	0.56	0.62	1.00	
<hr/>							
<i>Asia-Pacific</i>							
	BIS	AU	CN	ID	IN	JP	KR
BIS	1.00						
AU	0.61	1.00					
CN	0.21	0.25	1.00				
ID	0.10	0.17	0.29	1.00			
IN	0.56	0.56	0.22	0.14	1.00		
JP	0.43	0.51	0.20	0.25	0.37	1.00	
KR	0.50	0.51	0.30	0.40	0.32	0.49	1.00

in December 2015 central banks started to talk about “climate” and their dedication to the topic increased substantially from 2019 onward. Mentions seem to be higher after the September 2019 speech by Greta Thunberg at the UN Climate Action Summit in New York, and after the formal US withdrawal from the Paris Agreement in November 2020. (The United States re-entered in February 2021, before leaving again in January 2025). Similarly, the announcement by Facebook (later: Meta) of a proposed stablecoin called Libra (later: Diem) in June 2019 triggered central banks’ discourse on digital currencies. There were further mentions of these issues around key regulatory announcements, for instance the announcement of a blanket ban on crypto by the People’s Bank of China (PBOC) in

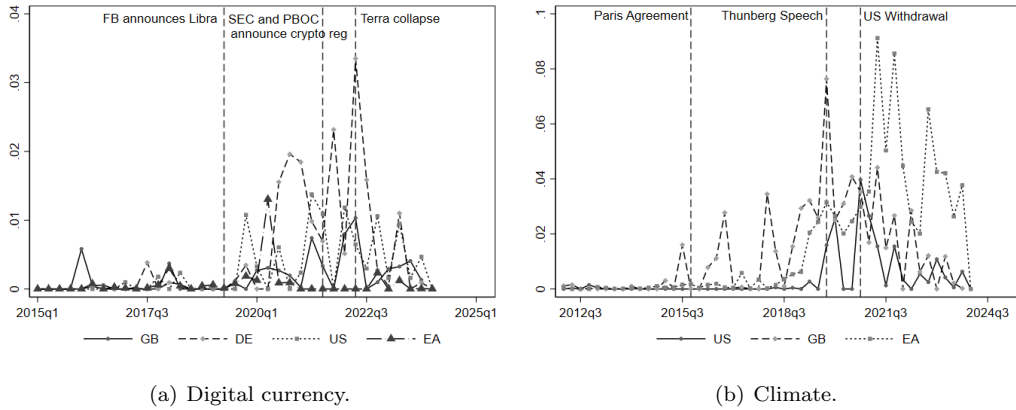
Figure 2: ATP of “monetary policy” and “inflation” over time.



Note: The lines show average document topic proportions by topic. The ATPs were averaged over all G-20 banks per quarter.

September 2021, as well as around the collapse of the Terra stablecoin in 2022.

Figure 3: Selected central banks’ ATPs for the topics “digital currency” and “climate” over time with key events.



Note: The lines show ATPs by topic and institution. The ATP were averaged per quarter.

4 Identifying thought leaders

4.1 Methodology

We propose that thought leaders are institutions whose usage of a specific topic in its speeches subsequently induces other institutions to increase their mentioning of the same topic. Conceptually, we are looking for the impact of the average topic proportion (ATP) of the thought leader (l) in one period $t - 1$ on the ATP of other institutions (i) in subsequent periods t . Since linear regression is not appropriate for bounded data and topic proportions are generally small, we model the $\log(ATP_{it})$ for country $i \in I$ as a function of all other countries at the previous time increment, $\log(ATP_{\ell,t-1})$, and the j^{th} macro-economic variables $X_{i,j,t-1}$ for each country and topic of interest allowing evolving impact of covariates on topic prevalence.

$$\log(ATP_{it}) = \alpha_i + \sum_{\ell \notin I} \log(ATP_{\ell,t-1})\beta_{i,\ell} + \sum_j \gamma_{i,j}X_{i,j,t} + \epsilon_{i,j,t} \quad (1)$$

Note that time is not explicitly included as a covariate, since the lagged institutions and economic variables are expected to capture time dependencies through more interpretable means. Since (1) is performed independently for each combination of country and topic of interest, we drop the subscript i to simplify notation.

To account for the macro-financial environment, we include a wide set of time-varying country-level statistics in $X_{i,j,t}$. In detail, we control for GDP growth, credit growth, inflation, the policy rate, an interbank rate where available, exchange rate movements against the US dollar, the evolution of the main stock index and its volatility, a financial condition index and the growth of the central bank's balance sheet.

The pre-eminent problem is to select the thought leaders while controlling for economic variables. Model selection is performed using lasso regression (Zou and Hastie, 2005) which explicitly sets some of the β_ℓ and γ_j coefficients to zero if they do not meaningfully improve the model fit as assessed by cross-validation. Statistically valid inference requires independent samples for building a model and constructing confidence intervals conditional

on the resulting model to avoid inflated type 1 error rates. However this relies on splitting the data into independent samples and incurring less efficient use of data. Alternatively, exact confidence intervals conditional on the model selection procedure are constructed using post-selection inference (PSI) using the full dataset (Kuchibhotla, Kolassa, and Kuffner, 2022). Unlike ordinary least squares (OLS) effect size intervals, PSI derives exact parameter confidence intervals and p-values accounting for the polyhedral constraints imposed by the model selection procedure.

Model (1) is fitted by optimizing (see Hastie, Tibshirani, and Wainwright (2015)):

$$[\hat{\beta}, \hat{\gamma}] = \arg \min_{\beta, \gamma} \left\{ \sum_t \left| \log(\mathbf{ATP}_t) - \left(\alpha + \sum_{\ell \notin I} \log(ATP_{\ell, t-1}) \beta_\ell + \sum_j \gamma_j X_{j, t} \right) \right|^2 + \lambda \|[\beta, \gamma]\| \right\}, \quad (2)$$

where λ is a penalty parameter. The appropriate λ value is selected to be the one whose associated model has the smallest root mean squared error. In the baseline, we use all available observations for every institution and topic. Beyond this global approach, we conjecture that institutions might be able to establish a leading role along two dimensions in the data: they could gain regional importance or topic-specific expertise. We therefore perform the optimisation detailed in equation 2 also by topic and by region where we include the BIS in all regions.

4.2 Results

We present the results from estimating equation 2 for selected topics in figures 4, 5 and 6. The colour-coded country code in the figure represents the lasso regression for that country. The detailed results from the regressions are in appendix B. Thus, we show the results from all 21 lasso regressions in one figure. The countries on the x-axis represent the potential leaders and the y-axis value indicates their impact on the colour-coded country. Sometimes observations were not sufficient to estimate the lasso regression; in those cases, the institution is not in the figure. Intuitively, thought leaders should accumulate many

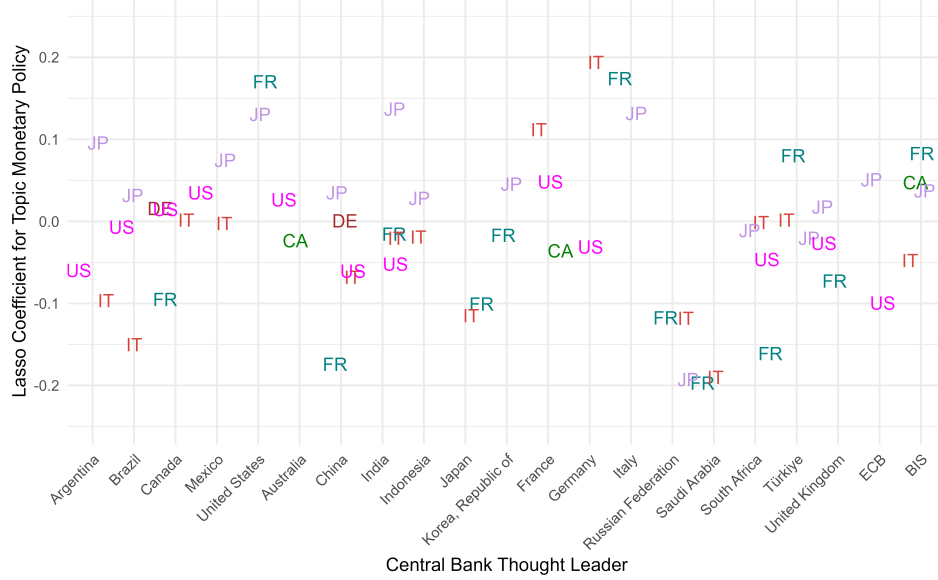
institutions' positive loadings and few or no negative ones.

First, we focus on monetary policy and inflation, two core topics for any monetary authority. For monetary policy, there are no obvious thought leaders. The results are depicted in the upper panel figure 4. Many coefficients hover around zero. And many institutions gather positive as well as negative loadings. For instance, officials from the Bank of France and Bank of Japan both increase their discussion of monetary policy following an increase from Bank of Italy the month prior. However, discussion of the topic by the Central Bank of Saudi Arabia, and to a lesser extent the People's Bank of China, actually decrease. For several institutions, the lasso did not pick up any foreign influence. These are the Bank of England, the South African Reserve Bank, Bank of Korea and Reserve Bank of India. (In the figure, for clarity, we only show the impact on G7 central banks. Doing so for the full G20 would not be legible. But the dataset is available upon request).

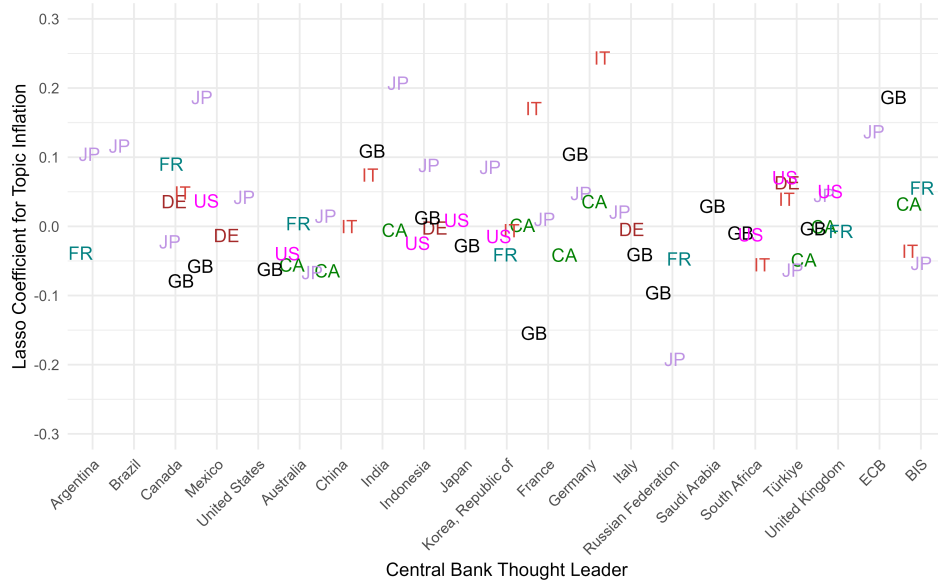
Based on our criteria, the US Federal Reserve could be considered a thought leader, since it is the only central bank without any significantly negative coefficients. Senior officials from both Bank of France and Bank of Japan tend to discuss monetary policy in speeches more frequently following discussion by officials from the Fed. However, the US is only selected for the models of these two central banks. Contrasting this with the inflation topic shown in the lower panel of figure 4, we find the Fed not to be influencing any other central bank regarding inflation. (Of course, the Fed may have a substantial influence on thinking by other central banks, but it may just not be measurable in topic proportions on the frequency of these two broad topics).

On closer look, regional leaders emerge. Both in Asia and Europe, we can identify leaders if we only consider within-region diffusion of topics. Japan (shown on the graph) and China (not shown) are following communications of other countries in the region, namely Korea and India, while Indonesia (not shown) also follows Korea. Therefore, when abstracting from non-Asian coefficients, India and Korea seem to be regional thought leaders. Interestingly, these are the two central banks that do not appear to follow the topics mentioned by others. Applying the same logic to the inflation topic, again Korea and India emerge as regional leaders. A similar regional pattern is present between France, Germany and Italy. These

Figure 4: Lasso coefficients across G7 institution models for monetary policy and inflation topic. The i^{th} country's two letter name code is plotted at with a vertical axis value of $\beta_{\ell \notin I}$ from (1) and aligned with country ℓ on the horizontal axis. Values are horizontally jittered to reduce overlay. The subset of models describing G7 institutions are selected here to increase clarity. The coefficients for all models can be found in Appendix B.



(a) Monetary policy.



(b) Inflation.

three central banks – which are part of the Eurosystem, together with the ECB – reference each other about monetary policy and receive negative coefficients from institutions outside Europe. Among these, Germany (Deutsche Bundesbank) follows topics in Canada (Bank of Canada) but no other central bank. However, the regional pattern is less pronounced for the inflation topic.

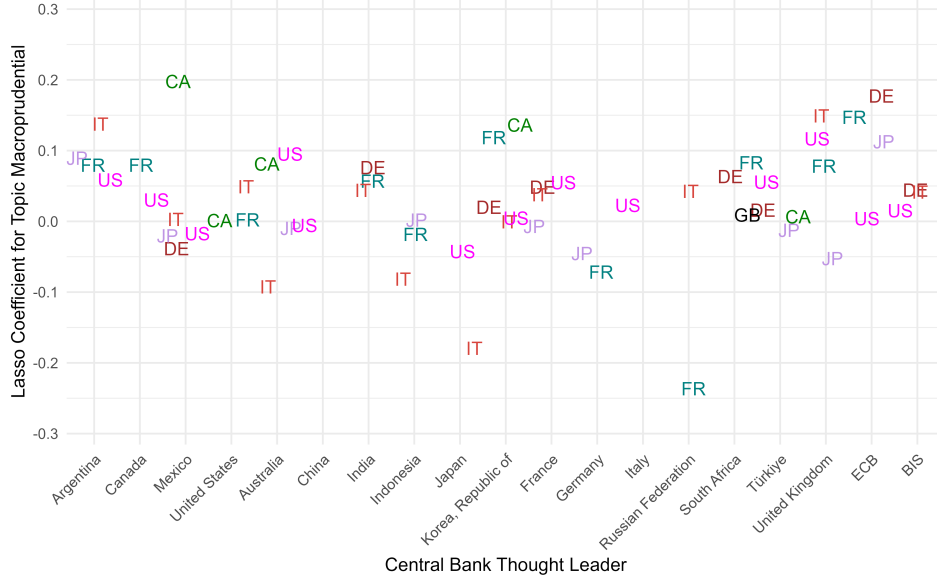
With regards to inflation, there are potential thought leaders in India and Germany; they are found to influence a number of central banks globally while also not selecting many banks in their own models. Also the ECB has a positive impact on mentions of the same topic by the Bank of England and Bank of Japan. For the covariates, many models select headline inflation, which is understandable, but with a low magnitude (see table B in the appendix).

Overall, these results emphasize that, regarding monetary policy and inflation, if there is influence of central banks on one another, it is primarily regional.

Next, we study speech mentions around central banks’ macroprudential policy work. Since macroprudential policy did not appear in that wording as a topic, we choose the broad topic “liquid”, whose word cloud reveals that this topic is associated with key words associated with central banks’ roles as regulators and lender of last resort (namely: stress, collateral, money market, financial_system, fund). The results are presented in figure 5.

In line with the results in [Park et al. \(2025\)](#), macroprudential policy seems to be a topic in which central banks are more connected to one another at a global level. Bank of England and Bank of Mexico are the only two institutions for which lasso picks fewer than three (in these cases: only one) other institutions. Our approach identifies the ECB and the central banks of Argentina, Korea and South Africa as thought leaders as they positively influence all models where they are selected. The ECB was selected in the models of Germany and France, which speaks to its regional mandate to serve and cooperate with the national central banks in the Euro area. The ECB was further selected by Japan, India and Australia and to a lesser degree the US, which might corroborate an international importance in the topic. Similarly, Argentina, Korea and South Africa were not just selected by their neighbours. However, the relative importance of Argentina has surprised us given the uniqueness of

Figure 5: Lasso coefficients across all institution models for macroprudential policy topic. The i^{th} country's two letter name code is plotted at with a vertical axis value of $\beta_{\ell \notin I}$ from (1) and aligned with country ℓ on the horizontal axis. Values are horizontally jittered to reduce overlay. The subset of models describing G7 institutions are selected here to increase clarity. The coefficients for all models can be found in Appendix B.



its macro-financial environment over the sample period. A possible explanation is that the topic might not only reflect central banks talking about potential risks but also facing actual crisis. We can further identify the Reserve Bank of India and the BIS as close to being thought leaders since they are also impacting a number of countries positively. Only speeches by officials of the Central Bank of Russia feature a negative impact.

Finally, we study the topic of payment systems, which represent another important central bank function and have gained significant prominence over the past years (see section 3.2). In the payment topic we see a number of thought leaders in Europe. Bank of Italy appears to be a regional thought leader, positively influencing France and the United Kingdom, and speeches by officials of the Bank of France show an association with those of central banks in Canada, Germany, Italy and the United Kingdom. The ECB partially adds to this trend, positively influencing Canada, Indonesia and the United Kingdom, but its discussion of this topic does not lead Argentina or Germany. In contrast, several central

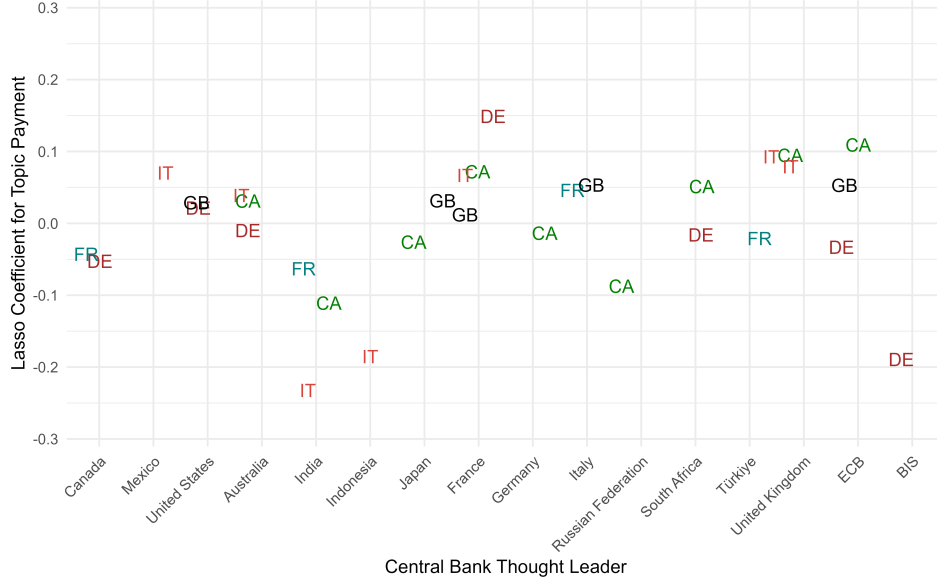
banks from Asia receive on average negative loadings. Also noteworthy is that several central banks from the Americas are missing, meaning they were discarded by all lasso models. This could be explained by the fact that some Latin American countries had low levels of financial access and digital payments over much of the sample period. However, lately these countries – particularly Brazil and Argentina – are also home to some of the most innovative payment companies, and have seen large strides in digital payment use. These innovations in the field of payments are not (yet) reflected in the topic of digital currencies. Overall, these results seem to imply that the payment topic only covers traditional functions of the central bank in providing and safeguarding the general infrastructure for national payments. More innovative themes in payments, such as digital currencies, fintech and fast payment systems may be reflected in other topics.

Note that data availability for each central bank, especially at the topic level, is not consistent. Banks with less communications per month have greatly reduced presence when analyzing a specific topic. However, lasso regression requires each covariate to have an observation at every time step that the modelled central bank does. Therefore, when working with the G20 central banks where data may not be available at consistent time steps, there are two options: remove time steps without complete data or impute the missing values. The first option is infeasible as this would result in too few time steps to fit the model. As such, imputation was chosen and any missing values or zeros were replaced with $1/40$. Unfortunately, with the G20 models, this can lead to the majority of covariates having various time steps imputed. This is of concern as the models may be fitting to this imputation rather than the observed values, if a large enough proportion of values are imputed. As such, the above discussion of findings is interesting, but will need further validation with additional methods.

4.2.1 Results considering only G7 Countries

As discussed earlier, the above results suffer from a lack of data, both due to countries with a minimal number of speeches in a topic and also short-lived topics across all countries. To mitigate this, a second analysis was conducted with just the countries within the G7

Figure 6: Lasso coefficients across all institution models for payments topic. The i^{th} country's two letter name code is plotted at with a vertical axis value of $\beta_{\ell \notin I}$ from (1) and aligned with country ℓ on the horizontal axis. Values are horizontally jittered to reduce overlay. The subset of models describing G7 institutions are selected here to increase clarity. The coefficients for all models can be found in Appendix B.



and four long-term topics: inflation, monetary policy, macroprudential and payments. This pruning of data allowed for each topic-country pair to contain at least 100 observations.

There are a few notable findings from this approach. First, when central banks are selected for a model, they usually have a smaller influence than the most, if not all, of macroeconomic covariates. This is seen both with the United Kingdom and the topic macroprudential (Figure 7 and Table 2) and with the United States and the topic monetary policy (Figure 8 and Table 3). For the United Kingdom and macroprudential, headline inflation and volatility are all selected in addition to Germany's discussion of macroprudential policy at a one month's lag. Volatility remains significant after post-selection inference, but headline inflation and Germany do not. Similarly, for the United States and monetary policy, headline inflation, policy rate, volatility, FCI and France at a one month lag are selected for the model. Here, only headline inflation remains significant after post-selection inference. This indicates that there may be influence from the lagged central banks, however they do

Figure 7: Lasso selection process for the topic 'macroprudential' for the United Kingdom. Each coloured line represents a potential predictor, with the vertical line representing the variables selected by the final model.

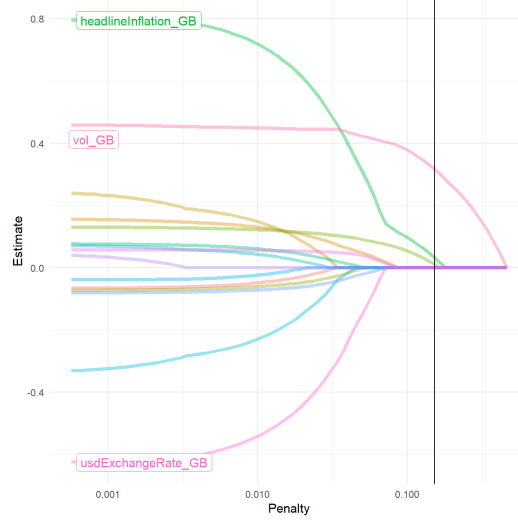


Table 2: Coefficients for the variables selecting in the lasso regression model for the United Kingdom with the topic 'macroprudential'

Predictor	Coefficient	Lower CI Bound	Upper CI Bound
Headline Inflation	0.221	-1.377	0.469
Volatility	0.496	0.104	1.089
Germany	0.138	-4.482	0.252

not maintain statistical significant after accounting for the model selection procedure. This is echoed with the model for Germany and monetary policy, where multiple banks are selected with coefficients comparable in magnitude to the macroeconomic variables. However, none remain statistically significant after post-selection inference (Figure 9 and Table 4). A full summary of the models can be found in Appendix B. But the majority of remaining bank-topic pairs exhibited low signal in comparison to noise and do not produce significant results.

Figure 8: Lasso selection process for the topic 'monetary policy' for the United States. Each coloured line represents a potential predictor, with the vertical line representing the variables selected by the final model.

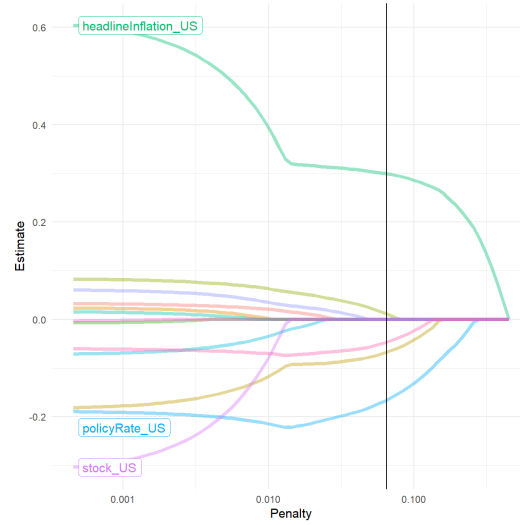


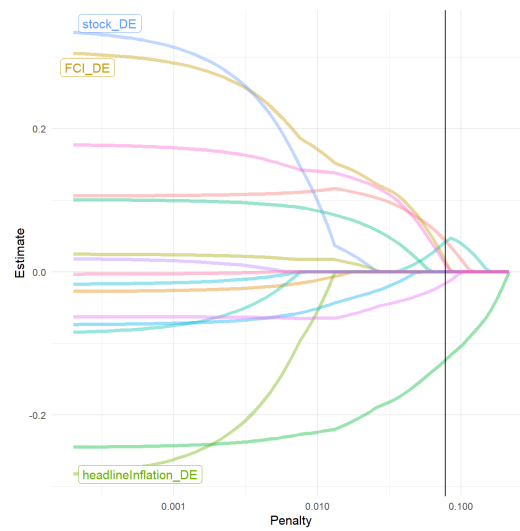
Table 3: Coefficients for the variables selecting in the lasso regression model for the United States with the topic 'monetary policy'

Predictor	Coefficient	Lower CI Bound	Upper CI Bound
Headline Inflation	0.319	0.167	0.550
Policy Rate	-0.230	-0.381	-0.055
Volatility	-0.089	-0.266	0.380
FCI	-0.111	-0.348	0.255
France	0.063	-1.511	0.157

Table 4: Coefficients for the variables selecting in the lasso regression model for Germany with the topic 'monetary policy'

Predictor	Coefficient	Lower CI Bound	Upper CI Bound
Policy Rate	-0.041	-2.245	4.836
USD Exchange Rate	0.164	-6.674	0.839
FCI	0.189	-7.309	1.655
Canada	0.135	-0.933	0.531
Italy	-0.208	-0.726	0.814
US	-0.065	-1.109	2.963

Figure 9: Lasso selection process for the topic 'monetary policy' for Germany. Each coloured line represents a potential predictor, with the vertical line representing the variables selected by the final model.



5 Conclusions

This paper examines how narratives on emerging and established policy topics diffuse across central banks, and whether some institutions systematically shape the communication of others. Using a large corpus of more than 11,000 speeches from senior officials of G20 central banks, the ECB and the BIS, we identified 40 topics through non-negative matrix factorisation and trace their evolution over time. We then combine these topic dynamics with macro-financial conditions and apply lasso regressions to detect institutions whose past communication influences the topical choices of their peers.

Three main conclusions emerge. First, central banks display persistent attention to core areas of their mandates — monetary policy, payments, reserve management and prudential issues — while also responding to major economic and financial developments. The topic model captures both cyclical patterns, such as the prominence of crisis-related themes during the GFC, and the rise of new issues including digital currencies, climate change and Covid-19. This confirms that public speeches provide a timely window into how central banks interpret and prioritise evolving challenges.

Second, we find systematic evidence that central banks pay attention to one another. While domestic conditions remain the dominant driver of monetary policy communication, especially for inflation and interest-rate topics, cross-border influence is visible in several other domains. Macroprudential policy discussions appear to be shaped by international dialogue, with the ECB emerging as a consistent thought leader. Similarly, European institutions play a central role in shaping narratives around payments. By contrast, the Federal Reserve — despite its global importance — does not consistently act as a thought leader across topics with this methodology.

Third, our results highlight that narrative diffusion is selective rather than universal. Some topics remain domestically anchored, while others spread across borders in ways that reflect institutional mandates, regional linkages and the global relevance of specific policy challenges. This suggests that central bank communication is both a domestic policy tool and a channel through which ideas and interpretations circulate internationally.

Of course, further work will be needed to assess the robustness of the results. Methodological choices can influence overall findings. If data quality permit it, there may be further insights when looking into a larger sample.

Taken together, these findings contribute to the understanding of how central banks form and transmit narratives in an interconnected world. They raise the possibility that communication is not only a mechanism for managing expectations at home, but also part of a broader international exchange of ideas that shapes how policymakers interpret economic and financial developments.

References

- Ahrens, M. and McMahon, M. Extracting economic signals from central bank speeches. In *Proceedings of the Third Workshop on Economics and Natural Language Processing*. Association for Computational Linguistics, 2021.
- Ahrens, M., Erdemlioglu, D., McMahon, M., Neely, C. J., and Yang, X. 2025. Mind your language: Market responses to central bank speeches. *Journal of Econometrics*, 249: 105921.
- Aizenman, J., Chinn, M. D., and Ito, H. 2016. Monetary policy spillovers and the trilemma in the new normal: Periphery country sensitivity to core country conditions. *Journal of International Money and Finance*, 68:298–330.
- Andersson, M., Dillén, H., and Sellin, P. 2006. Monetary policy signaling and movements in the term structure of interest rates. *Journal of Monetary Economics*, 53(8):1815–1855.
- Armeliuss, H., Bertsch, C., Hull, I., and Zhang, X. 2020. Spread the word: International spillovers from central bank communication. *Journal of International Money and Finance*, 103:102116.
- Arseneau, D. M. and Osada, M. 2023. Central bank mandates and communication about climate change evidence from a large dataset of central bank speeches. Technical report, Bank of Japan.
- Bertsch, C., Hull, I., Lumsdaine, R. L., and Zhang, X. 2024. Four facts about international central bank communication. Technical report, Sveriges Riksbank Working Paper Series.
- Bholat, D., Broughton, N., Ter Meer, J., and Walczak, E. 2019. Enhancing central bank communications using simple and relatable information. *Journal of Monetary Economics*, 108:1–15.
- Binder, C. 2017. Fed speak on main street: Central bank communication and household expectations. *Journal of Macroeconomics*, 52:238–251.

- Blei, D. M. 2012. Probabilistic Topic Models. 55(4). ISSN 0001-0782.
- Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., and Jansen, D.-J. 2008. Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4):910–945.
- Born, B., Ehrmann, M., and Fratzscher, M. 2014. Central bank communication on financial stability. *The Economic Journal*, 124(577):701–734.
- Bruno, V. and Shin, H. S. 2015. Capital flows and the risk-taking channel of monetary policy. *Journal of Monetary Economics*, 71:119–132.
- Campbell, J. R., Evans, C. L., Fisher, J. D., Justiniano, A., Calomiris, C. W., and Woodford, M. 2012. Macroeconomic effects of federal reserve forward guidance [with comments and discussion]. *Brookings papers on economic activity*, pages 1–80.
- Conrad, C. and Lamla, M. J. 2010. The high-frequency response of the eur-usd exchange rate to ecb communication. *Journal of Money, Credit and Banking*, 42(7):1391–1417.
- Davis, J. S. and Wynne, M. A. 2019. Central bank communications. *The Oxford Handbook of the Economics of Central Banking*, page 263.
- Du, B., Fung, S., and Loveland, R. 2018. The informational role of options markets: Evidence from fomc announcements. *Journal of Banking & Finance*, 92:237–256.
- Ehrmann, M. and Talmi, J. 2020. Starting from a blank page? semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, 111:48–62.
- Feldkircher, M. and Teliha, V. 2024. Speeches in the green: The political discourse of green central banking. *Energy Economics*, 135:107646.
- Ferrara, F. M. and Angino, S. 2022. Does clarity make central banks more engaging? lessons from ecb communications. *European Journal of Political Economy*, 74:102146.
- Glas, A. and Müller, L. S. 2021. Talking in a language that everyone can understand? transparency of speeches by the ecb executive board. Technical report, Working Paper.

- Hansen, S., McMahon, M., and Prat, A. 2018. Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics*, 133(2):801–870.
- Hansen, S., McMahon, M., and Tong, M. 2019. The long-run information effect of central bank communication. *Journal of Monetary Economics*, 108:185–202.
- Hansson, M. 2021. Evolution of topics in central bank speech communication. arXiv preprint arXiv:2109.10058.
- Hastie, T., Tibshirani, R., and Wainwright, M. 2015. Statistical learning with sparsity. *Monographs on statistics and applied probability*, 143(143):8.
- Hwang, I. D., Lustenberger, T., and Rossi, E. 2023. Central bank communication and public trust: The case of ecb speeches. *Journal of International Money and Finance*, 137: 102916.
- Kryvtsov, O. and Petersen, L. 2021. Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics*, 117:760–780.
- Kuchibhotla, A. K., Kolassa, J. E., and Kuffner, T. A. Mar. 2022. Post-Selection Inference. *Annual Review of Statistics and Its Application*, 9(1). ISSN 2326-8298, 2326-831X. doi: 10.1146/annurev-statistics-100421-044639. URL <https://www.annualreviews.org/doi/10.1146/annurev-statistics-100421-044639>.
- Larsen, V. H. and Thorsrud, L. A. 2018. Business cycle narratives. Technical report, Norges Bank Working Paper No 3/2018.
- Miranda-Agrippino, S. and Rey, H. 2020. Us monetary policy and the global financial cycle. *The Review of Economic Studies*, 87(6):2754–2776.
- Park, T., Perez-Cruz, F., and Shin, H. S. 2025. Mapping the space of central bankers’ ideas. BIS Working Paper, (1299).

- Petropoulos, A. and Siakoulis, V. 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(4):141–153.
- Phelan, G. and Campbell, D. A. 2026. Testing hypotheses of covariate effects on topics of discourse. *Statistical Analysis and Data Mining: An ASA Data Science Journal*, in press. doi: 10.1002/sam.70066.
- Roos, M. and Reccius, M. 2024. Narratives in economics. *Journal of Economic Surveys*, 38(2):303–341.
- Shiller, R. J. 2017. Narrative economics. *American Economic Review*, 107(4):967–1004.
- Siklos, P., Amand, S. S., and Wajda, J. 2018. The evolving scope and content of central bank speeches. Technical report, CIGI Papers No. 202 November 2018.
- Tibshirani, R. J., Taylor, J., Lockhart, R., and Tibshirani, R. 2016. Exact Post-Selection Inference for Sequential Regression Procedures. *Journal of the American Statistical Association*, 111(514):600–614.
- Warin, T. and Stojkov, A. 2023. “decoding” policy perspectives: Structural topic modeling of european central bankers’ speeches. *Journal of Risk and Financial Management*, 16(7): 329.
- Wu, X., Nguyen, T., and Luu, A. T. Jan. 2024. A survey on neural topic models: methods, applications, and challenges. *Artificial Intelligence Review*, 57(2):18. ISSN 1573-7462. doi: 10.1007/s10462-023-10661-7.
- Zou, H. and Hastie, T. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320.

A Appendix 1: LDA stability

Table 5: LDA is unstable with respect to the random seed. The table shows the topic numbers into which the top 5 terms are placed in running LDA with 10 topics initialized with 4 different random seeds. Rows are for different topics, columns are for seeds.

	seed 100	seed 101	seed 110
topic 1	econom, economi, inflat, rate, will	econom, economi, increas, inflat, will	econom, economi, growth, increas, will
topic 2	also, bank, develop, payment, will	bank, market, not, risk, will	also, bank, develop, payment, will
topic 3	area, central_bank-, euro, not, will	also, bank, develop, payment, will	area, countri, euro, not, will
topic 4	bank, market, not, risk, will	can, inflat, not, on, polici	bank, market, not, risk, will
topic 5	can, economi, growth, not, on	area, central_bank-, countri, euro, not	can, inflat, monetary_policy-, not, polici

B Appendix 2: Lasso Regression Results

Table 6: Lasso coefficients for topic monetary policy. Institution for which the lasso regression was done is indicated in the column headers. Missing coefficients indicate that the variable was not selected by lasso.

	AR	AU	CA	CN	FR	DE	IN	ID	IT	JP	KR	MX	RU	SA	TR	GB	US
<i>Potential thought leaders</i>																	
Argentina	.	.	.	.	.	.	.	.	-0.10	0.10	.	.	.	.	.	.	-0.06
Australia	.	.	-0.02	.	.	.	.	.	.	.	.	.	.	.	.	.	0.03
BIS	-0.02	.	0.05	.	0.08	.	.	.	-0.05	0.04	.	.	-0.06	.	-0.07	.	.
Brazil	.	.	.	.	-0.15	.	.	.	-0.15	0.03	.	.	-0.18	.	.	.	-0.01
Canada	.	.	.	.	-0.10	0.02	.	.	0.00	.	.	.	.	.	.	.	0.01
China	-0.50	.	.	.	-0.17	0.00	.	.	-0.07	0.03	.	.	.	.	.	.	-0.06
France	.	.	-0.04	.	.	.	.	.	0.11	.	.	.	.	.	.	.	0.05
Germany	-0.01	-0.00	.	.	-0.02	.	.	.	0.19	.	.	-0.07	.	.	.	.	-0.03
India	0.52	-0.02	.	0.09	-0.02	.	.	.	-0.02	0.14	.	-0.02	.	.	.	.	-0.05
Indonesia	.	.	.	.	.	.	.	.	-0.02	0.03	.	.	.	.	.	.	.
Italy	.	-0.02	.	.	0.17	.	.	.	-0.12	0.13	.	-0.12	.	.	.	.	.
Japan	.	-0.04	.	.	-0.10	.	.	0.05	-0.12	.	.	.	.	.	.	.	.
Korea Rep.	.	-0.00	.	0.17	-0.02	.	.	-0.10	-0.00	0.07	.	.	.	.	.	.	0.03
Mexico	.	.	.	.	.	.	.	.	-0.12	-0.19	.	.	.	.	.	.	.
Russian Fed.	.	.	.	.	-0.12	.	.	.	-0.12	.	.	.	.	.	.	.	.
Saudi Arabia	.	.	.	.	-0.20	.	.	.	-0.19	.	.	.	.	.	.	.	.
South Africa	.	-0.01	.	.	-0.16	.	.	.	-0.00	-0.01	.	0.14	.	.	0.02	.	-0.05
Türkiye	-0.08	.	.	.	0.08	.	.	.	0.00	-0.02	.	.	.	.	.	.	.
UK	.	.	.	.	-0.07	.	.	.	.	0.02	.	.	.	.	.	.	-0.03
US	.	.	.	.	0.17	.	.	.	.	0.13	.	.	.	.	.	.	.
<i>Macro-financial covariates</i>																	
Credit Growth	.	.	.	.	.	.	.	.	-0.08	.	.	.	0.00	.	.	.	-0.02
ECB	.	.	.	.	.	.	.	.	.	0.05	.	.	.	.	.	.	-0.10
FCI	.	-0.13	-0.16	.	.	.	.	.	.	.	.	.	.	.	0.12	.	-0.15
Headline Inflation	-0.01	0.01	-0.01	0.03	-0.06	.	-0.02	0.03	0.03	-0.04	.	-0.04	0.02	.	0.04	-0.00	0.05
Policy Rate	0.04	.	0.10	.	.	.	.	.	.	.	.	.	.	.	.	.	-0.05
QE & QT	-0.00	-0.00	-0.00	.	0.00	-0.00	-0.01	0.02	-0.01	-0.02	-0.02	-0.06	.	-0.00	0.01	-0.00	0.01
Real GDP	.	.	.	.	.	.	0.01	-0.00	0.06	-0.02	.	.	.	.	.	0.02	.
Stock Mkt. Index	.	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00
USD Exch. Rate	0.08	.	.	.	.	.	-0.03	0.00	.	-0.03	-0.00	.	-0.06	.	.	.	.
Volatility	.	.	.	.	-0.65	.	.	.	.	.	.	.	.	.	.	.	-0.05

Table 7: Lasso coefficients for topic payment. Institution for which the lasso regression was done is indicated in the column headers. Missing coefficients indicate that the variable was not selected by lasso.

	AR	AU	CA	CN	FR	DE	IN	ID	IT	JP	KR	MX	RU	SA	TR	GB	US
<i>Potential thought leaders</i>																	
Argentina	.	.	.	.	.	.	.	.	-0.10	0.10	.	.	.	.	.	.	-0.06
Australia	.	.	0.03	.	.	-0.01	0.01	-0.14	0.04	0.04	0.06	.	.	.	.	.	0.03
BIS	.	.	.	.	-0.19	-0.19	.	0.28	-0.05	0.04	0.27	.	-0.06	.	-0.07	.	.
Brazil	.	.	.	.	-0.15	.	.	.	-0.15	0.03	.	.	-0.18	.	.	.	-0.01
Canada	.	0.07	.	.	-0.04	-0.05	.	-0.01	0.00	0.03	-0.02	.	-0.09	.	.	.	0.01
China	.	.	.	.	.	.	.	.	-0.07	0.03	.	.	.	.	.	.	-0.06
France	.	.	0.07	.	.	0.15	.	-0.07	0.11	.	.	.	.	.	.	0.01	0.05
Germany	.	.	-0.01	.	.	.	.	0.23	0.19	0.14	-0.09	-0.07	.	.	.	.	-0.03
India	.	.	-0.11	0.09	-0.06	.	.	-0.01	-0.02	0.13	.	-0.02	.	.	.	.	-0.05
Indonesia	.	.	.	.	.	.	.	.	-0.02	0.03	.	.	.	.	.	.	.
Italy	.	-0.02	.	.	0.05	.	.	-0.14	-0.12	0.13	.	-0.12	.	.	.	0.05	.
Japan	.	-0.04	-0.03	.	-0.10	.	0.00	-0.12	-0.12	.	.	.	.	.	.	0.03	.
Korea Rep.	.	-0.00	.	0.17	-0.02	.	.	-0.24	.	0.05	.	.	0.13	.	.	.	.
Mexico	.	.	.	.	.	.	.	-0.07	-0.00	0.07	.	.	.	.	.	.	0.03
Russian Fed.	.	.	-0.09	.	.	.	.	0.11	-0.12	-0.19	.	.	.	.	.	.	.
Saudi Arabia	.	.	.	.	-0.20	.	.	.	-0.19	.	.	.	.	.	.	.	.
South Africa	.	-0.01	0.05	.	-0.16	-0.02	-0.01	0.55	-0.00	-0.01	0.02	0.14	.	.	0.02	.	-0.05
Türkiye	.	.	.	.	0.08	.	.	-0.40	0.00	-0.02	.	.	-0.31	.	.	.	-0.03
UK	-0.14	.	0.09	.	.	.	.	.	0.08	0.02	.	.	.	.	.	.	.
US	.	.	.	.	0.02	.	.	0.51	.	0.13	-0.08	.	-0.12	.	.	0.03	.
<i>Macro-financial covariates</i>																	
Credit Growth	-0.20	.	.	.	.	.	.	.	-0.08	.	.	.	0.00	.	.	.	-0.02
ECB	.	.	0.11	.	.	-0.03	.	0.07	.	0.05	.	.	.	.	.	0.05	-0.10
FCI	.	-0.13	-0.16	.	.	.	.	.	.	.	.	.	.	.	0.12	.	-0.15
Headline Inflation	0.00	0.05	0.02	-0.01	0.05	.	0.01	.	0.03	-0.04	-0.03	-0.04	0.02	.	0.04	-0.00	0.05
Policy Rate	-0.10	.	-0.15	.	.	.	.	-0.55	.	.	.	.	0.04	.	.	.	-0.05
QE & QT	-0.00	0.00	-0.00	-0.04	-0.01	0.01	-0.01	-0.02	-0.01	-0.02	-0.02	-0.06	.	-0.00	0.01	-0.00	0.01
Real GDP	.	.	-0.12	.	.	.	0.01	.	0.06	-0.02	.	.	.	.	.	0.02	.
Stock Mkt. Index	.	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00
USD Exch. Rate	0.08	.	.	.	.	.	-0.03	0.00	.	-0.04	-0.00	.	-0.06	.	.	.	.
Volatility	.	.	.	.	-0.65	.	.	0.30	.	.	.	.	.	.	.	.	-0.05

Table 8: Lasso coefficients for topic inflation with G7 countries alone. Institution for which the lasso regression was done is indicated in the column headers. Missing coefficients indicate that the variable was not selected by lasso and * indicates statistical significance.

	CA	FR	DE	IT	JP	GB	US
<i>Potential thought leaders</i>							
Canada	-	0.471	0.166	-0.117	0.008	.	-0.004
France	-0.215	-	0.380	0.355	-0.076	-0.314	-0.074
Germany	0.094	0.055	-	0.362	0.202	.	-0.031
Italy	.	0.023	-0.266	-	-0.066	.	-0.021
Japan	.	-0.365	0.135	-0.104	-	.	-0.015
UK	.	-0.183	0.202	-0.139	0.054	-	0.107
US	.	0.078	.	0.300	-0.054	.	-
<i>Macro-financial covariates</i>							
Credit Growth	.	-0.053	-0.190	-0.299	-0.236	.	-0.091
FCI	-0.614*	0.193	0.151	-0.619	-1.223*	.	-0.148
Headline Inflation	.	0.356	-0.078	0.355	0.261	.	-0.065
Policy Rate	0.631*	0.201	0.455	0.802	0.076	.	0.071
QE & QT	.	-0.106	-0.027	0.074	0.042	.	-
Real GDP	.	-0.088	.	-0.197	0.073	.	0.139
Stock Mkt. Index	0.298	-0.176	0.344	-0.464	-1.374	.	0.438
USD Exch. Rate	.	-0.174	0.476	0.096	0.303	-0.388	-
Volatility	.	-0.454	0.262	0.150	-0.002	.	0.048

C Appendix 3: Topic definitions

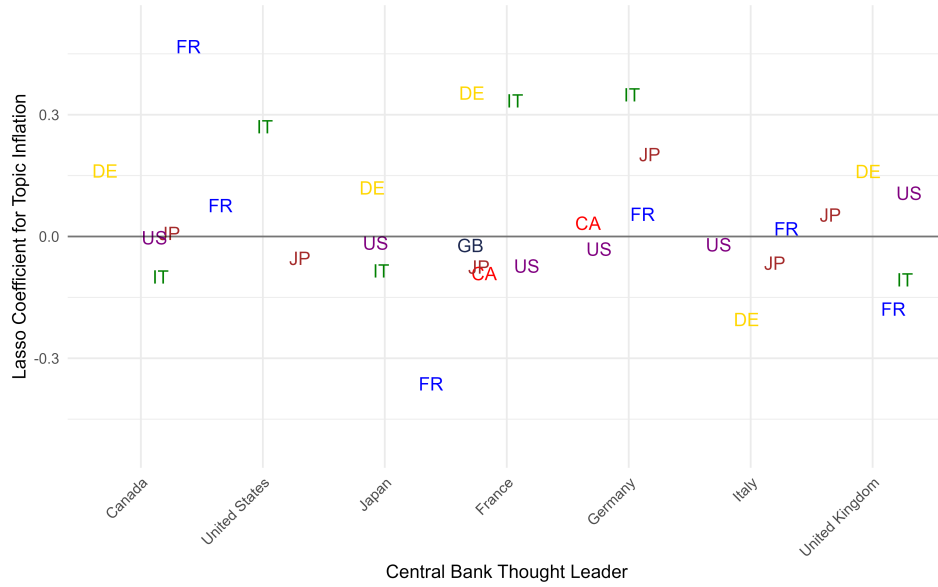
Table 9: Lasso coefficients for topic macroprudential with G7 countries alone. Institution for which the lasso regression was done is indicated in the column headers. Missing coefficients indicate that the variable was not selected by lasso and * indicates statistical significance.

	CA	FR	DE	IT	JP	GB	US
<i>Potential thought leaders</i>							
Canada	-	0.226	.	0.140	0.013	.	0.056
France	.	-	.	0.052	-0.136	.	0.215*
Germany	.	-0.229	-	-0.242	-0.364	0.138	-0.066
Italy	.	0.212	.	-	0.262	.	0.172
Japan	.	.	0.098	-0.511*	-	.	-0.241*
UK	.	0.140	.	0.215	-0.115	-	0.239*
US	0.147	0.224	0.300*	0.225	0.156	.	-
<i>Macro-financial covariates</i>							
Credit Growth	0.202	.	.	0.168	-0.044	.	0.100
FCI	0.183	0.269	.	0.403	0.616	.	0.058
Headline Inflation	.	-0.396	-0.241	0.085	-0.727	0.221	0.823*
Policy Rate	.	.	.	.	-0.175	.	-0.176
QE & QT	0.242	.	.	0.419	0.106	.	-
Real GDP	.	.	.	0.110	-0.278	.	-0.060
Stock Mkt. Index	.	0.003	.	0.202	0.046	.	-0.957*
USD Exch. Rate	-0.372	-0.242	-0.205	-0.269	0.405	.	-
Volatility	0.247	.	0.208	-0.025	0.120	0.496	-0.198

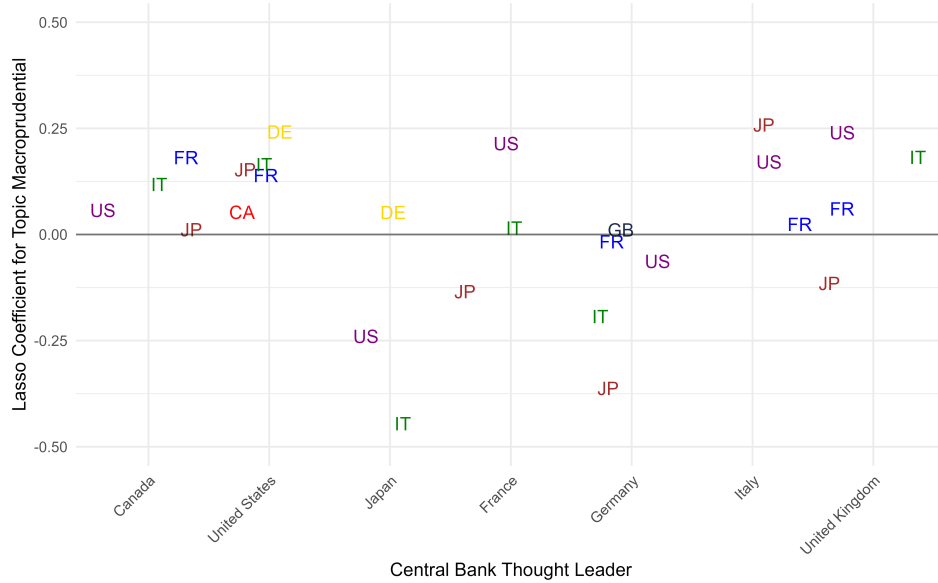
Table 10: Lasso coefficients for topic monetary policy with G7 countries alone. Institution for which the lasso regression was done is indicated in the column headers. Missing coefficients indicate that the variable was not selected by lasso and * indicates statistical significance.

	CA	FR	DE	IT	JP	GB	US
<i>Potential thought leaders</i>							
Canada	-	0.032	0.135	0.196	-0.037	-0.294*	.
France	-0.141	-	.	0.250	-0.091	-0.046	0.063
Germany	.	0.137	-	0.174	0.043	-0.125	.
Italy	-0.074	0.357	-0.208	-	0.314*	-0.169	.
Japan	.	-0.253	.	-0.328	-	0.075	.
UK	.	-0.174	.	0.067	0.095	-	.
US	.	0.101	-0.065	0.210	0.327*	.	-
<i>Macro-financial covariates</i>							
Credit Growth	.	0.087	.	-0.293	-0.099	-0.135	.
FCI	-0.266	-0.083	0.189	-0.574	.	0.335	-0.111
Headline Inflation	.	-0.302	.	0.606	.	0.216	0.319*
Policy Rate	0.289	0.244	-0.041	0.606	0.111	0.065	-0.230*
QE & QT	.	0.096	.	-0.009	-0.270*	-0.123	-
Real GDP	.	.	.	0.091	.	.	.
Stock Mkt. Index	.	.	.	-0.280	0.370*	.	.
USD Exch. Rate	0.048	.	0.164	-0.277	-0.607*	.	-
Volatility	.	-0.444	.	0.261	0.142	0.084	-0.089

Figure 10: Lasso coefficients across reduced G7 institution models for topics inflation and macroprudential. The i^{th} country's two letter name code is plotted at with a vertical axis value of $\beta_{\ell \notin I}$ from (1) and aligned with country ℓ on the horizontal axis. Values are horizontally jittered to reduce overlay.



(a) Inflation.

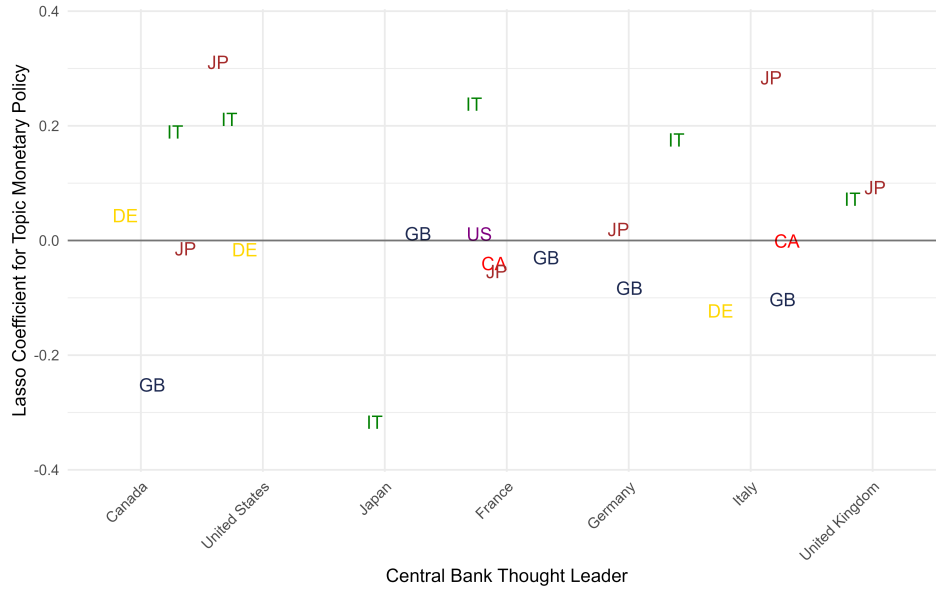


(b) Macroprudential.

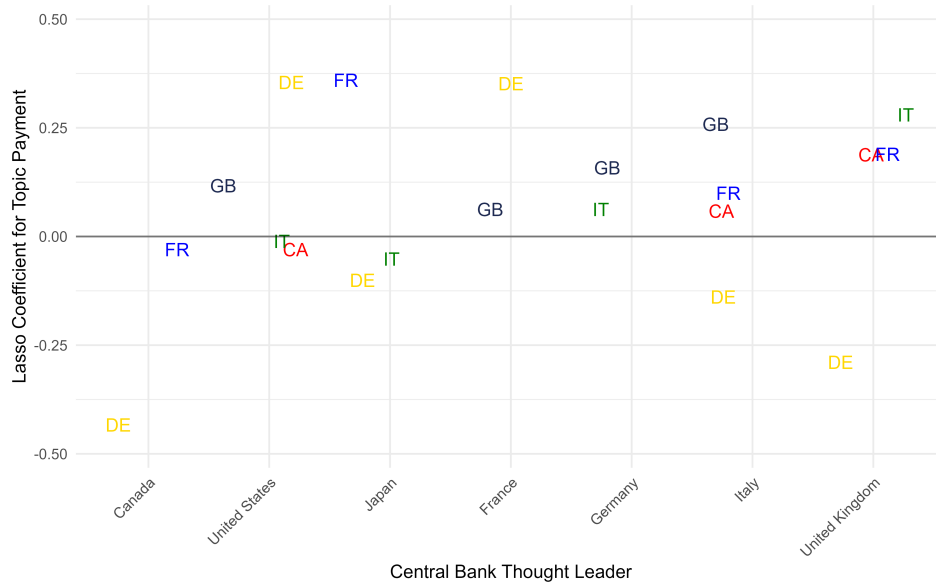
Table 11: Lasso coefficients for topic payment with G7 countries alone. Institution for which the lasso regression was done is indicated in the column headers. Missing coefficients indicate that the variable was not selected by lasso and * indicates statistical significance.

	CA	FR	DE	IT	JP	GB	US
<i>Potential thought leaders</i>							
Canada	-	-0.243	-0.435*	.	-0.097	.	-0.148
France	.	-	0.351	.	-0.240	0.107	0.005
Germany	.	.	-	0.050	0.232	0.232	0.032
Italy	0.121	0.333	-0.141	-	-0.068	0.324	0.132
Japan	.	0.520	-0.101	-0.054	-	.	0.138
UK	0.272	0.254	-0.292	0.312	0.156	-	-0.025
US	-0.140	.	0.354	-0.034	-0.171	0.261	-
<i>Macro-financial covariates</i>							
Credit Growth	.	-0.106	-0.173	0.057	.	.	0.125
FCI	0.109	-0.328	0.423	1.818*	.	.	0.164
Headline Inflation	.	0.788*	-0.922	0.278	.	.	-0.481
Policy Rate	-0.175	.	-0.233	-1.462	.	.	-0.076
QE & QT	0.161	-0.148	-0.110	0.240	0.152	-0.261	-
Real GDP	-0.247	.	0.206	-0.185	0.101	.	0.165
Stock Mkt. Index	.	0.481	1.836*	1.942	-0.178	0.370	0.371
USD Exch. Rate	0.210	.	0.616	.	.	0.447	-
Volatility	.	0.747	0.545*	0.484	-0.209	.	0.063

Figure 11: Lasso coefficients across reduced G7 institution models for topics monetary policy, and payment. The i^{th} country's two letter name code is plotted at with a vertical axis value of $\beta_{\ell \notin I}$ from (1) and aligned with country ℓ on the horizontal axis. Values are horizontally jittered to reduce overlay.

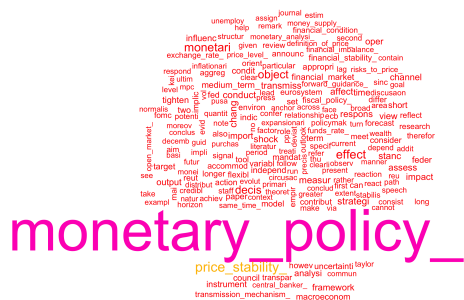


(a) Monetary Policy.

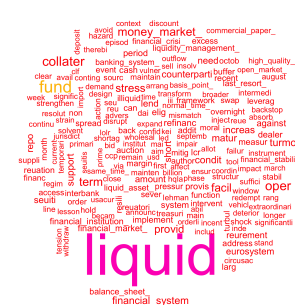


(b) Payment.

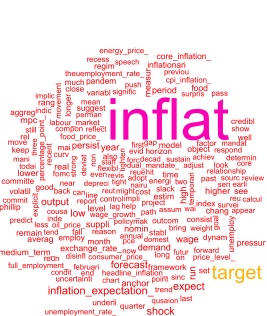
Figure 12: Wordclouds of the top 200 terms within each of the topics identified using the full dataset including G20, ECB, and BIS speeches.



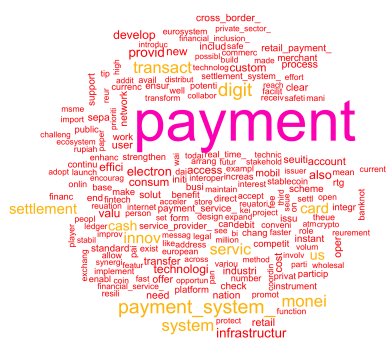
(a) Monetary Policy.



(b) Macroprudential Policy.

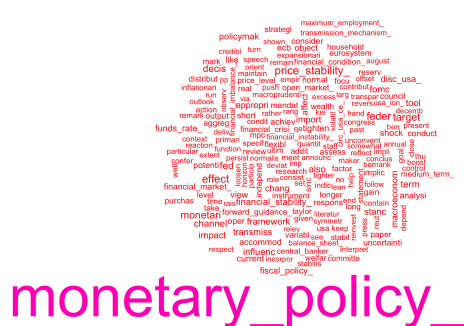


(c) Inflation.



(d) Payment.

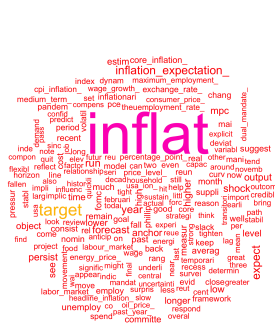
Figure 13: Wordclouds of the top 200 terms within each of the topics identified using the full dataset including G7 speeches.



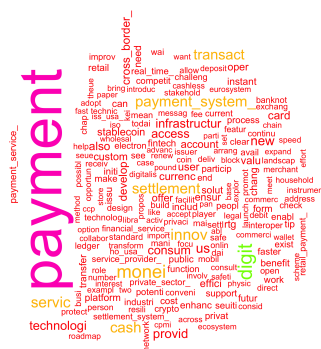
(a) Monetary Policy.



(b) Macroprudential Policy.



(c) Inflation.



(d) Payment.