// NO.26-005 | 02/2026

DISCUSSION PAPER

// FRIEDRICH HEINEMANN AND JAN KEMPER

ECB Council Members' Objectives and Public Debt at Home – Evidence From an AI-Based Textual Analysis

Leibniz Association

ZEW

ECB Council Members' Objectives and Public Debt at Home – Evidence from an AI-Based Textual Analysis

Friedrich Heinemann (ZEW Mannheim and University of Heidelberg)

Jan Kemper (ZEW Mannheim and University of Mannheim)

February 2026

Abstract:

We examine the changing attention that ECB Governing Council members pay to different policy objectives by analysing more than 4,600 speeches given between the establishment of the ECB and the summer of 2024. Alongside the primary objective of price stability, we consider the following potential secondary objectives: financial stability, stability of the government bond market, sustainable public debt, climate protection and distribution. On the methodological side, we take advantage of LLMs to identify the speeches' coverage of each of these objectives and the associated support. We conduct a series of validation tests to verify our AI-based scores, including a conventional dictionary approach. We use two-way fixed effects regressions to search for a link between a country's level of public debt and the objective function of its representatives. The results suggest that objectives have become more diverse in recent years. An increase in the public debt-to-GDP ratio in a governor's home country is associated with a shift in focus away from the primary objective and towards a growing coverage and support for secondary objectives. This general pattern is particularly robust for the distribution objective. These results can only be partly explained by governor selection. Therefore, in their communication, individual governors indicate shifts in their objective function in response to changes in the fiscal situation of their home country.

JEL Classification: E58, E52, H63

Keywords: fiscal dominance, green monetary policy, large language model, text analysis

Acknowledgment: We gratefully acknowledge financial support from the Stiftung Geld und Währung.

1 Introduction

At a general level, the objective of the ECB is well defined in the EU Treaties. According to Art. 127 (1) TFEU, the primary objective is to maintain price stability. Without prejudice to this, the ECB is required to support the general economic policies of the EU in order to achieve the objectives of the Union. While this is precise in terms of the priority given to price stability, it leaves a wide scope for the evolution of the objective function as a whole.

Clearly, since the introduction of the euro, monetary policy debates have pointed to a broadening of the objective function as a consequence of shocks, newly emerging societal challenges, possible learning effects and shifts in the EU's overall policy objectives. The years of the financial crisis have shown that financial stability deserves considerable attention from central banks. Moreover, the ECB has taken over the crucial responsibility for banking supervision (Howarth and Quaglia, 2014), which shifts financial and banking stability directly into the central bank's mandate. Subsequently, during the euro area debt crisis and pandemic, the ECB increasingly established itself as a lender of emergency liquidity to governments through various substantial sovereign bond purchase programmes (Havlik and Heinemann, 2021; Wyplosz, 2019), expressing the ECB's growing attention to public debt developments and the stability of sovereign bond markets. Issues relating to climate and distribution are less directly related to the monetary policy domain, but they definitely have received an increasing weight in the general EU policy agenda. Thus, they represent further dimensions that could play a role in the ECB's objective function in relation to the Treaty requirement to support the general EU objectives (on green monetary policy: Schnabel, 2021; on the responsibility of monetary policy for social inclusion: Dossche et al., 2021).

This increasing range of potential second-order objectives beyond the priority of price stability raises difficult questions for the ECB and has sparked controversy. The growing involvement of the Eurosystem as the largest investor in the euro area sovereign bond market has fuelled the debate on the risks of "fiscal dominance" and an increasing central bank attention for the liquidity of highly indebted euro countries, with similar discussions for other central banks engaging in asset purchases (Bordo and Levy, 2021; Cochrane, 2023; Hinterlang and Hollmayr, 2020; Sargent and Wallace, 1981; Schnabel, 2020; Wyplosz, 2019). The concern is that the ECB may compromise price stability and be reluctant to fight inflationary risks in a sufficiently forward-looking manner due to the possible negative impact of monetary tightening on sovereign bond markets of highly indebted euro countries. With regard to various other societal objectives, such as climate protection and social inclusion, critics point to the risks of a broader range of policy objectives for the legitimacy of the ECB's independence (Kronberger Kreis, 2021; Wyplosz, 2019). The more monetary policy decisions are guided by societal objectives beyond price stability, the more central bank independence may come under threat (De Haan and Eijffinger, 2016).

Our contribution does not engage with this normative debate. Instead, we pose two positive questions that inform these controversies. First, we ask how we can empirically measure

and describe the monetary policy objectives of ECB officials, both across individuals and over time. Second, we investigate whether there is an empirical pattern in which the level of public debt in the home country is related to the type of objective function revealed by members of the Governing Council. Specifically, we investigate whether governors from high-debt countries pay more attention to secondary ECB objectives beyond the primary concern of maintaining price stability. We exploit the panel structure of our data to determine the extent to which changes in debt levels are associated with changes in the policy objectives of a country's representatives in the ECB Council. We also consider whether these changes reflect evolving individual views or a selection mechanism for a country's governors. Under the latter, member states would appoint a different type of governor in response to changes in the national fiscal situation.

The euro area is an ideal setting in which to study the relationship between fiscal conditions and monetary policy objectives. It combines a centralized monetary policy with substantial fiscal autonomy at the national level, creating pronounced cross-country heterogeneity in fiscal conditions within a single currency union. This makes it possible to study the link between heterogenous fiscal circumstances within a common currency union and regional representatives' positions in the governing council. Unlike single-country time-series analyses or simple cross-country comparisons, this distinct setup enables us to isolate the effects of fiscal conditions on monetary policy preferences while keeping the monetary policy environment constant.

With this interest, we contribute first to the literature on the impact of elected politicians on monetary policy decision making and its interplay with central bank independence (Belke and Potrafke, 2012; Cahan et al., 2019; Demiralp et al., 2019; Ehrmann and Fratzscher, 2011; Giesenow and De Haan, 2019; Masciandaro, 2022). For example, Giesenow and de Haan (2019) reject the hypothesis that the ideology of a national government has an impact on monetary policy decisions of the country's independent central bank. Instead of considering the partisan dimension of politics as done in this literature, we focus on political interests related to the country's fiscal situation.

Second, we add to the literature that uses text analytical approaches to identify individual preferences of monetary policymakers that has evolved over more than two decades (since e.g. Chappell Jr et al., 1997). In this literature, governing council members' preferences are measured on the public transcripts of council meetings or recordings of public speeches (Apel et al., 2022; Apel and Grimaldi, 2014; Bjerkander and Glas, 2024; Bohl et al., 2023; Fadda et al., 2025; Gardner et al., 2022; Hansen et al., 2018; Shapiro and Wilson, 2022). This text-analytical approach is of particular importance for ECB-related research because roll-call votes are not published. In the literature, differences between the preferences of Council members on the 'hawk-dove' dimension have been analysed in particular (Bennani and Neuenkirch, 2017; Heinemann and Kemper, 2021; Malmendier et al., 2021; Picault and Renault, 2017). In most cases, dictionary approaches are used in which words are defined that describe a certain aspect (e.g. inflation risk).

This existing literature shows that individual characteristics, background and inflation experience have an influence on the preferences of monetary policymakers (Bennani and Neuenkirch, 2017; Chappell Jr et al., 2008; Malmendier et al., 2021). Kokoszczyński and Mackiewicz-Łyziak (2024) show for members of the Polish Central Bank Council that previous political activities make dovish positioning more likely.

We introduce two innovations to the existing text-as-data literature on central bank communication: one conceptual and one methodological. Conceptually, we are the first to search for a multidimensional objective function for central bankers that goes beyond the traditional binary hawkishness/dovishness dimension, which only identifies stronger or weaker concerns about price stability. Instead, we explore six possible dimensions of monetary policy objectives. This allows us to determine the extent to which tendencies towards a more blurred, multidimensional objective function exist, and how this trend differs between central bankers from different countries. On the methodological side, we take advantage of the recently improved access to large language models (LLM). We use GPT-4o mini and GPT-5 mini to search a large corpus of 4,680 central bank speeches given by ECB board members and national governors over 25 years between the late founding phase of the euro in 1997 and 2024. This allows us to identify the changing importance of different objectives over time and across individuals (e.g. board members versus governors). By comparing the findings of traditional lexicographic textual analysis with those of LLM, we also contribute to assessing the robustness and reliability of the new AI tools for monetary policy research based on text-as-data.

Our analysis focuses on the following six dimensions that might play a role in a central bank's objective function:

- *Price stability*: This objective refers to the ECB's primary objective as it is defined in Art. 127 TFEU. Measuring the individual support for this objective has already received substantial attention in the literature on monetary policy hawks and doves.
- *Financial stability*: The second objective we include is financial stability. The financial crisis 2008/09 had brought the early imperfections of European banking and financial market regulation to the forefront. Subsequently, the ECB had to accept major responsibility for banking supervision with its leading role in the Single Supervisory Mechanism and its direct supervision of significant banks since 2014 (ECB, 2014). Hence there are good arguments to include financial stability as one possible dimension of a central bank objective function.
- *Sustainable public debt*: Central bankers discuss issues relating to public debt and deficits, for example, in the context of debates on fiscal policy and the appropriate macroeconomic policy mix. They may also refer to the preferred level and dynamics of public debt when discussing debt sustainability risks relevant to monetary policy, as these risks can affect the real economy or financial stability.
- *Sovereign bond markets*: A somewhat different dimension of fiscal issues relates to the market for sovereign bonds. Sovereign bond market tensions with rising interest rate

spreads have been at the centre of the euro area debt crisis in 2010/12. In the pandemic, ECB representatives have increasingly emphasized that rising sovereign spreads may impede the monetary transmission mechanism. More recently, the ECB has even established a permanent "Transmission Protection Instrument" (TPI) that, under certain conditions, wants to counteract "a deterioration in financing conditions" (European Central Bank, 2022), subject to eligibility criteria like the compliance with the EU fiscal framework. The ECB considers increases in spreads that are not aligned with country-specific fundamentals to be a potential obstacle to the monetary transmission process. As discussed above, critics would see an increasing ECB attention for the financing conditions of member states as inappropriate and as a step towards fiscal dominance (for a recent debate: Boffa and Treeck, 2025).

- *Climate protection*: In recent years, the debate on the role of the ECB in tackling climate change seems to have received growing attention (Christine Lagarde, 2021). The ECB also employs increasingly more resources to cover climate policy topics. For instance, in 2021, a climate change centre within the ECB was established that directly reports to Christine Lagarde (Christine Lagarde, 2021). The inclusion of a possible climate policy objective into our testing design is hence well justified.
- *Distribution*: Monetary policy, through its impact on growth, employment, inflation and (real) asset returns, can have distributional consequences. Central banks may pay more or less attention to these repercussions. Like climate policy, distributive objectives of monetary policy could enter the ECB objective function through the requirement to support the general objectives of the Union.

Without delving further into the debates about what a central bank should do, this paper measures the changing emphasis on the primary objective and the various secondary objectives, revealing some cross-country and cross-individual patterns behind these changes.

The findings indicate a growing diversity in objectives over recent years, with heightened emphasis on the ECB's role in climate policy and distributional issues in the years before the pandemic and the subsequent inflation hike. Notably, significant variation exists across countries and individuals in the ECB Council. Central bank governors from nations with higher debt-to-GDP ratios are more likely to address topics such as climate change, public debt, government bond markets, and distributional concerns, even when accounting for euro area macroeconomic conditions through year fixed effects.

Controlling for country and year fixed effects, there is a robust correlation between the debt-to-GDP ratios and a reduced emphasis on price stability among national central bank governors, accompanied by a more dovish policy stance. Among the secondary objectives considered, the increased focus on distributional issues is particularly robust. These findings suggest that rising public debt in a governor's home country is associated with a reallocation of attention away from price stability toward secondary policy objectives. Selection in the appointment of central bankers explains part, but not all, of the observed effect.

In the next section, we describe the corpus of ECB speeches that is the foundation for our analysis. In Section 3, we explain our use of LLMs, its validation and the dictionary robustness check and, in Section 4, we present our results on the prominence of the above defined objectives over time and across countries. Section 5 presents our panel regression results exploring the link between national public debt and the individual central bankers' emphasis on the various objectives. Section 6 explores further the role of governor selection for our findings and possible non-linearities, followed by conclusions.

2 Data and descriptives

In order to detect differences and changes in the topic salience of speeches given by ECB Council members, a large speech corpus is required. The most extensive source available is the central bank speech collection from the Bank of International Settlements (BIS) (Bank for International Settlements, 2024). The BIS uploads speeches from almost all national central banks worldwide on a daily basis, dating back to 1997. The database includes speeches from the executive board and from the national governors in the ECB governing council. All euro member governors are covered by speeches but there are significant differences in the frequencies, which will be discussed in detail below.

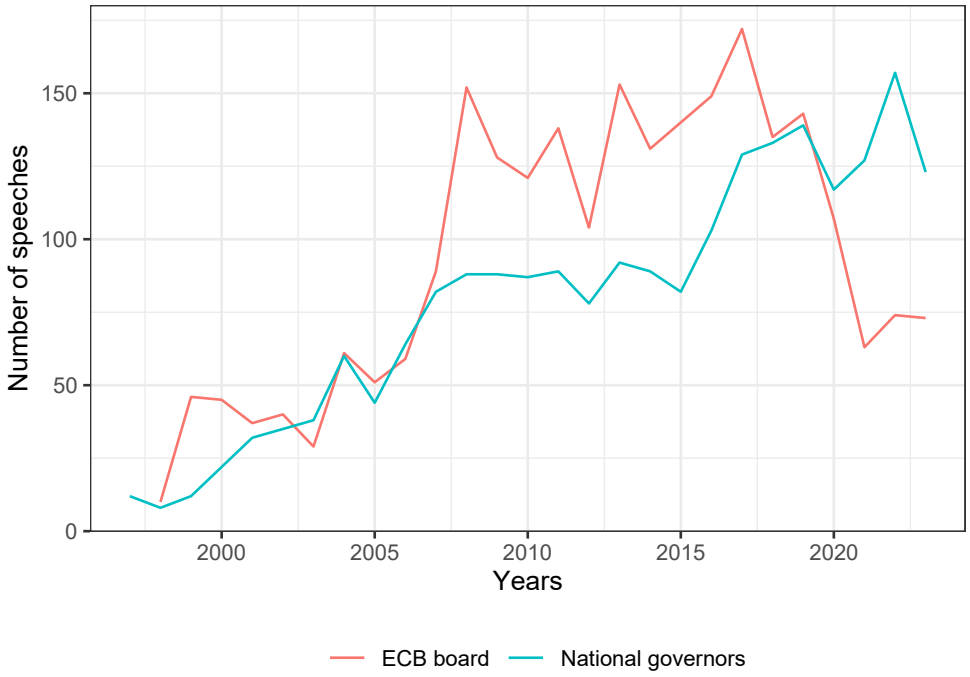
The analysis is restricted to English speeches or speeches with an English transcript because we are interested in the euro area discourse. Furthermore, English speeches are better captured by the BIS website, and the analysis is more comparable across states. The main caveat of the BIS database is that it does not cover the whole universe of speeches. There are two sources of missing data. First, speeches are only available if an official transcript exists. Second, some of the speeches are not detected by the BIS, hence they are missing in the database. To estimate the magnitude of the problem, we compare the speeches from German governors uploaded on the Bundesbank website with the BIS sample. For Germany's former national governor Jens Weidmann, for example, 216 transcripted speeches are available, while the BIS data base reports 161. Considering that several speeches on the Bundesbank website are only in German and thus not included in the BIS data base, the BIS sample seems to be a good approximation of the total number of speeches given to an international audience. Moreover, there is no indication that the BIS curates speeches based on their content. We therefore consider the BIS sample to be broadly representative of the full set of speeches.

A series of pre-processing steps were taken to enable statistical analysis of the text data. The final data set entails 4,680 speeches and covers the period from February 1997 until

June 2024.[1] 1,142 speeches were held by executive board members (excluding presidents and vice presidents), 997 are speeches by the presidents, 345 are from the vice presidents and 2,196 originate from the national governors. We restrict the sample exclusively to national governors and members of the executive board (including president and vice president) during their terms in office.

In Figure 1, the number of published speeches by year and position is plotted. The red line depicts ECB board member speeches, and the blue line displays the yearly number of speeches by national governors. Overall, the number of speeches increased over time, which also reflects the enlargement of the euro area. From 2006 until 2008 there was a steep rise which coincides with the beginning of the financial crisis. Afterwards the number of given speeches remained high in comparison to the early phase of the euro. After 2015, there was another steep increase in speeches for both national governors and ECB board members, but this trend only continued for the governors. During the Corona crisis fewer speeches were given by ECB board members. On the other side, this downward trend was less pronounced for national governors.

Figure 1: Number of speeches per year

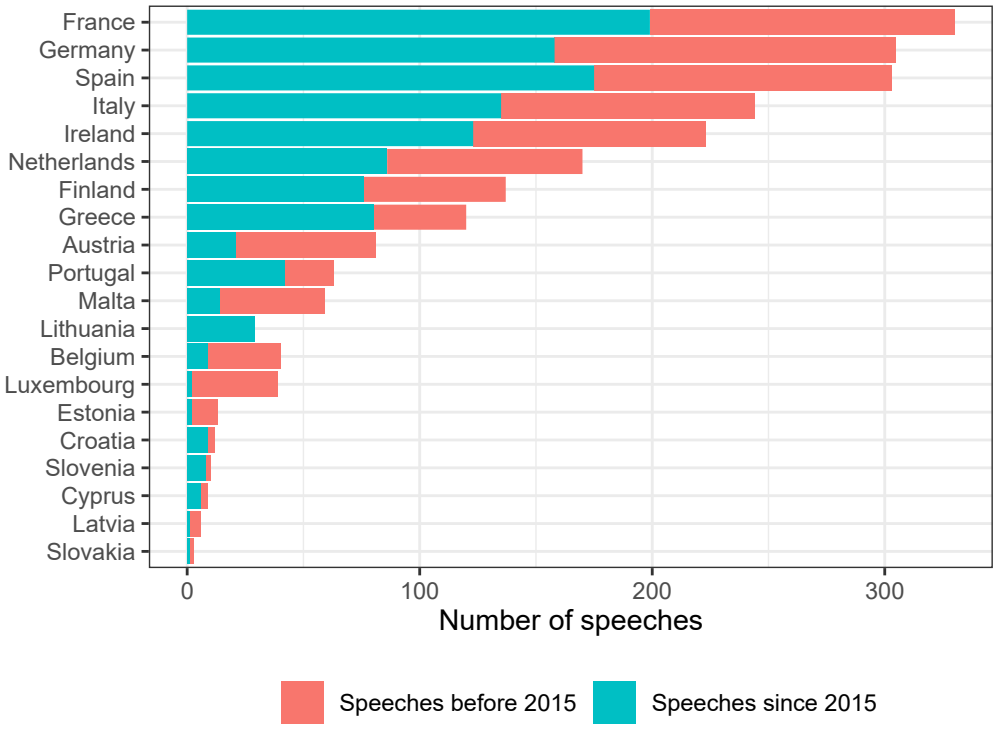


Notes: The figure is based on data from the BIS. Speeches from 2024 are excluded from the graph as the sample ends mid-2024.

[1] The ECB was founded in 1998. The BIS database starts at 1997, we also included the 13 available speeches in the late founding phase of the ECB in 1997.

The distribution of speeches is not uniform across countries. In Figure 2, the recorded speeches of the national governors are plotted by country. The red bar shows the total number of speeches held before 2015 while the blue bar indicates the number of speeches that were released since 2015. Not surprisingly, country size matters and national governors from France, Germany, Spain, and Italy are the most vocal ones whereas the smaller Eastern European countries have the lowest number of published speeches. The correlation between country size and the number of speeches is also robust when just speeches of the last eight years are included which guarantees that all countries (except Croatia) had already adopted the euro over the full period under analysis. Therefore, we must bear in mind that results based on this corpus of speeches are more heavily influenced by observations from larger euro area countries.

Figure 2: Number of speeches from national governors by country



Notes: The plot is based on data from the BIS. The x-axis shows the number of speeches, the colours indicate which speeches were held before 2015.

3 Scoring methods

3.1 Method choice

We identify what objectives central bankers discuss in these speeches and how much emphasis they assign to these objectives by applying text-analytical methods. Assuming that central bankers talk more about objectives that they consider important, we could use the frequencies to infer how relevant a certain objective is in the central bank discourse, how this has changed over time and how the emphasis varies across countries and individual council members. As described in the introduction, we concentrate on the following six possible explicit or implicit objectives which play a relevant role in the monetary policy discourse: “price stability”, “financial stability”, “public debt”, “sovereign bond markets”, “climate change” and “distribution”.

Qualitatively reading 4,680 speeches and classifying the text into topics is resource intensive. We apply computational text analysis methods which are designed to (semi-) automatize the analysis. For identifying the salience of the objectives and measure positions in speeches from the ECB Council members, we apply two distinct methods. First, as our baseline, we apply a large language model (GPT-4o mini for topic analysis and GPT-5 mini for opinion scoring) to measure the salience and positioning of central bankers towards these objectives. Second, as a robustness check, we develop and use a new dictionary which measures the explicit and implicit objectives of the ECB across our six dimensions.

Both methods have distinct advantages and disadvantages. LLMs have great language capabilities and can detect nuances in human language where dictionaries classically fail.[2] But this comes at a cost because LLMs are black boxes raising issues of transparency. Dictionaries which have lower language capabilities are only a rough measure but they are transparent and results are easily replicable for third parties. To increase transparency, we, therefore, provide robustness checks based on the lexicographic approach and validate our LLM classifications on manually annotated data.

3.2 Large Language Model

With the LLMs, we aim to measure both the prevalence of the objectives in the speeches and the position of the speaker towards the respective objective. The prevalence score measures the mere coverage of an objective, regardless of the speaker’s position on it. The

[2] For example, a dictionary which should measure the topic “climate change” could detect that the following sentence is about climate change: “Reducing climate change is a responsibility of a central bank”. An LLM can go far beyond that and also recognize the direction of the statement. Here it would also find that the topic is “climate change” but additionally it concludes that the speaker demands more engagement of central banks in climate protection.

position score captures the degree to which the speaker expresses support for the relevance and importance of a particular objective in monetary policy decision-making.

We apply GPT-4o mini for the prevalence analysis and GPT-5 mini for the position score. In a series of qualitative and quantitative tests, these models performed best in terms of classification quality, speed and costs (see Appendix A.1.2. for more information on the model selection). For each of our six categories, we have a specific (but analogous) prompt which implies that each speech is submitted to the model six times.

For the prevalence score, we first classify how frequent a certain topic is mentioned on a continuous scale from 1.0 to 5.0. A prevalence score of 1.0 reflects the absence of any reference to the objective, whereas a score of 5.0 indicates that the speech is entirely devoted to that objective.

In the second part of the prompt, we ask for the speaker's position toward the respective objective across a continuous scale from -1 to 1 (position score). If the speaker strongly expresses the view that the central bank should be more involved in achieving the objective, the position score is -1.0. If the position is neutral, the score is 0.0. A score of 1.0 is given if the speaker firmly believes that this objective should not influence monetary policy decisions and should not distract from the objective of price stability. Therefore, the position score shows how much emphasis a central banker places on the primary objective compared to each secondary objective. A larger value represents a stronger emphasis on price stability relative to the alternative objective.

For example, the prompt for the climate protection objective is in the box below (all prompts are fully documented in Appendix, Table A.3). First, we indicate the role of the language model as an “economist” to give a context to the model.[3] Then, the prompt asks if the topic climate change or climate protection is discussed anywhere in the speech and to assign the topic prevalence score between 1.0 and 5.0. Finally, a definition of the position dimension is given and the model should classify the speech along the continuous position scale from -1.0 to +1,0. -1.0 represents the speaker’s strong support for the increased use of monetary policy tools in climate politics, while 1.0 represents the opposite view, that the central bank should focus less on climate policy.

[3] The context is given to the model in the initial „system role“ section of the prompt.

Prompt climate protection objective

> You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech of a member of the European Central Bank Council.
>
> To what extent does the following speech pay attention to the topic climate change or climate protection? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic).
>
> In a second step, classify the opinion of the speaker towards the role of central banks for climate policy. Return a continuous indicator value with one decimal place from -1.0 to +1.0 with the following interpretation for the extreme and the central indicator values: -1.0 is appropriate if the speaker expresses strongly that central banks should use and adjust their instruments to play a more active role in climate policy in the future; 0.0 is appropriate if the speaker expresses the opinion that central banks already support climate policy exactly right; 1.0 is appropriate if the speaker thinks strongly that central banks should focus less on climate policy and concentrate instead more on the traditional objectives of monetary policy like price stability.
>
> The output format should only be: ["Answer 1", "Answer 2"]. For Answer 2, return NA if climate change or climate protection is not discussed at all.

The prompts for the other secondary objectives are constructed in an analogous way. For price stability, the position score is exactly the hawkish-dovishness dimension. -1.0 would describe a speech in which the speaker does not see any inflationary risk and supports a more expansionary use of monetary policy instruments and 1.0 would stand for the opposite case of high inflation concerns and the support for a restrictive monetary policy stance.

The main underlying assumption is that the GPT models understand the categories and the speeches sufficiently to return reasonable classifications. In different studies on central bank texts, LLMs already resulted in promising classifications (Woodhouse and Charlesworth, 2023). However, we provide below several validation checks to assess our AI-based scores.

3.3 Composite secondary objectives prevalence index

In addition to the single objective scores, we also make use of a composite prevalence score that is synthesized from the LLM outputs for the single objectives. To get a meaningful summary statistic which describes the focus on the primary mandate of price stability in comparison to all other here analysed secondary objectives, we develop the following “composite secondary objectives prevalence index”:

$$Y_i = \frac{Price\ Stability\ Prevalence\ Score_i\ - \frac{1}{k}\sum_{j=1}^{k} Secondary\ Objectives\ Prevalence\ Score_{ji}}{4}$$

The composite index subtracts the price stability prevalence score by the unweighted mean of the prevalence scores of all other five categories. It is normalized by four which ensures a range between minus one to plus one. A value of -1,0 indicates that the speech Y_i is entirely about all secondary objectives and does not cover price stability at all. A value of +1,0 indicates that the speech is entirely about price stability and does not cover any of the here selected secondary objectives.

3.4 LLM validation

To ensure that our LLM model adequately captures the intended concepts, we employ three complementary validation strategies. First, we look at the resulting LLM scores in the context of the macroeconomic environment. We assess whether the observed trajectories reflect contemporaneous economic and crisis-related developments in a plausible way. Second, we benchmark the model outputs against manually annotated data. Both validation approaches support the use of the topic prevalence and position score. Third, for topic prevalence, we compare our LLM scores with conventional dictionary scores (sections 3.5 and 5.3.2).

Section 4 below presents the descriptive time-series results for both measures across all categories. Overall, the time series in Figures 3 and 5 are consistent with contemporaneous macroeconomic developments and related policy debates within the ECB. For instance, in 2022, when inflation in the euro area surged, ECB Council members increasingly emphasized price stability and adopted a more hawkish stance. Similarly, during the euro area sovereign debt crisis, fiscal objectives received heightened attention, while concerns about financial stability rose sharply at the onset of the global financial crisis. This qualitative evidence supports the hypothesis that the LLMs provide meaningful scores.

To further assess the validity of the LLM outputs, we trained human research assistants to manually annotate a subset of the speeches. The resulting validation dataset is then compared with the model's predictions (see Appendix section A.1.3). To increase precision and comparability, we segmented the speeches into paragraphs consisting of five sentences each. This choice followed experimentation with paragraphs of three and ten sentences; five-sentence segments strike a balance between providing sufficient contextual information and allowing for the construction of a sufficiently large validation sample. We assume that these five-sentence segments are representative of the classification task applied to full speeches. From this pool, we randomly selected 600 paragraphs, which were

manually classified by research assistants across all categories. The first 50 annotations were cross-checked jointly with the authors, revealing a high level of agreement.

Table 1 reports the main results of the human–model comparison across all objectives. Column 1 presents the mean absolute error (MAE) for the topic prevalence score, defined as the average absolute difference between human annotations and model predictions. Given that the score ranges from 1 to 5, MAE values between 0.09 and 0.95 indicate a high level of agreement. Overall, these results suggest close alignment between human and model assessments.

Column 2 reports an approximate accuracy measure. Because accuracy is defined for categorical variables, we classify a prediction as correct if it lies within a ±1 window of the human label. Under this definition, agreement is high across all categories, with the lowest value observed for financial stability (0.73). Given substantial class imbalance in smaller categories, both MAE and accuracy are imperfect performance indicators. We therefore report the raw evaluation data and linear correlations in Figure A.1 in the appendix, which reveal a medium to strong correlation between human and model labels, explaining a large share of the observed variation.

For the topic position score, MAEs are higher but still indicate that the model captures meaningful variation. The score ranges from −1 to 1, with an average MAE of 0.44 across categories. Accuracy within a ±0.5 window ranges from 0.46 to 0.70. Appendix Figure A.2 shows that model and human labels are positively correlated for all categories, with particularly strong alignment for price stability and weaker correlations for financial stability, sovereign bond markets, and distribution. This pattern is expected, as position classification is more complex and likely constrained by the limited context provided by five-sentence excerpts.

Overall, the validation exercise demonstrates that the LLM captures substantial and economically relevant information, supporting its use as a measurement tool in this analysis.

Table 1: Validation LLMs

	Topic prevalence		Position score	
Objective	**MAE**	**Accuracy**	**MAE**	**Accuracy**
Price Stability	0.48	0.90	0.41	0.68
Financial Stability	0.95	0.73	0.48	0.54
Public Debt	0.35	0.91	0.46	0.59
Sovereign bond markets	0.50	0.91	0.49	0.46
Distribution	0.46	0.93	0.42	0.67
Climate protection	0.09	0.99	0.36	0.70

Note: MAE is the abbreviation for Mean Absolute Error (MAE) which is the mean of the absolute value deviation between a human coder and the applied model predictions. The accuracy for topic prevalence is calculated by accepting a window of ± 1 between human and model values as an agreement. Accuracy is the share of the correct prediction of the total predictions. For the accuracy of the position score a window ± 0.5 is allowed.

3.5 Dictionary approach as further validation check

LLM-based textual scorings have quickly entered economic research and the method is about to establish itself firmly in the toolbox of empirical research (Hassan et al., 2025). However, LLM approaches are confronted with the problem of limited transparency due to the highly complex algorithms and their non-deterministic nature. To address these transparency issues, we develop a more conventional dictionary approach to measure the prevalence of the objectives as a further validation check.

A dictionary includes a list of words that belong to specific concepts or topics which they should measure. As all terms are pre-defined this method is highly transparent and replicable. We limit the robustness check to the prevalence score where a dictionary approach can be expected to provide reasonable results since this score is only about the coverage of a topic, whereas a dictionary approach can hardly offer an informative robustness check for the position score. On positions, an LLM approach is inherently more powerful since the position score wants to quantify a text's assessment of an objective and, hence, requires understanding of the full textual context.

In our dictionary approach, all term frequencies of a respective category are counted and then summed up on the speech level. The speech length between different speakers varies by a lot, therefore, the frequencies are normalized by the number of words per speech. The following formula describes the normalization process:

$$Topic\ prevalence\ score\ dictionary_{sc} = \frac{Dictionary\ terms_{sc}}{Total\ terms_{s}}$$

s indicates a speech and *c* stands for a dictionary category. The topic prevalence score for this robustness check is just the frequency of the counted dictionary terms of category *c* in speech *s* divided by the total terms in speech *s.* In other words, it is the share of dictionary terms in a given speech.

A dictionary approach requires a high quality and context specific dictionary which correctly measures the intended concept. To ensure valid results, we develop our dictionary in a multi-stage semi-supervised procedure combining machine learning methods and qualitative expertise (for an exact description see Appendix A.1). Table A.1 in Appendix section A.1 presents the resulting dictionary.

4 Objectives over time and per country

4.1 Objectives over time

In this section, we present the resulting patterns of prevalence and position scores over time. For prevalence, we present both LLM and dictionary scores. This enables us to assess the proximity of the two scores and, consequently, the reliability of the LLM approach.

Topic prevalence

In the first analysis, the saliency of the six explicit or implicit objectives over time is examined. Figure 3 shows the LLM results for topic prevalence. The black line shows the price stability category which is an important topic throughout. In the early 2000s the price stability prevalence was on a particularly high level. During this period, the inflation rate was above the two percent target rate of the ECB (World Bank, 2022). In 2009, the prevalence of price stability declined while the discourse on financial stability was on the rise. This result fits very well with the events unfolding during that time. From 2007 until 2009 the global financial crisis hit the world and at the same time, the discourse on financial stability became more prominent. In 2009, the inflation declined close to zero (World Bank, 2022). During the following years, price stability remained a salient topic although less dominant than in the early 2000s. In the year 2015, the inflation rate reached the zero percent threshold and hence a region with a deflation risk; this coincides with an increase in central bankers' public discourse on price stability. In 2022, the price stability category steeply rose reflecting the dramatic hike in inflation which found a strong resonance in central bankers' communication.

For the other categories, there are three main findings. First, the financial stability category became more prevalent in the discourse after the outbreak of the financial crisis. Before the crisis, financial stability was already regularly mentioned in central bankers' speeches but on average the level almost doubled after the outbreak of the crisis. The financial crisis and the euro crisis seem to have shaped the discourse of the ECB Council which fits to the institutional changes emerging from the lessons of the crisis with the ECB taking over major responsibility in the emerging European Banking Union.

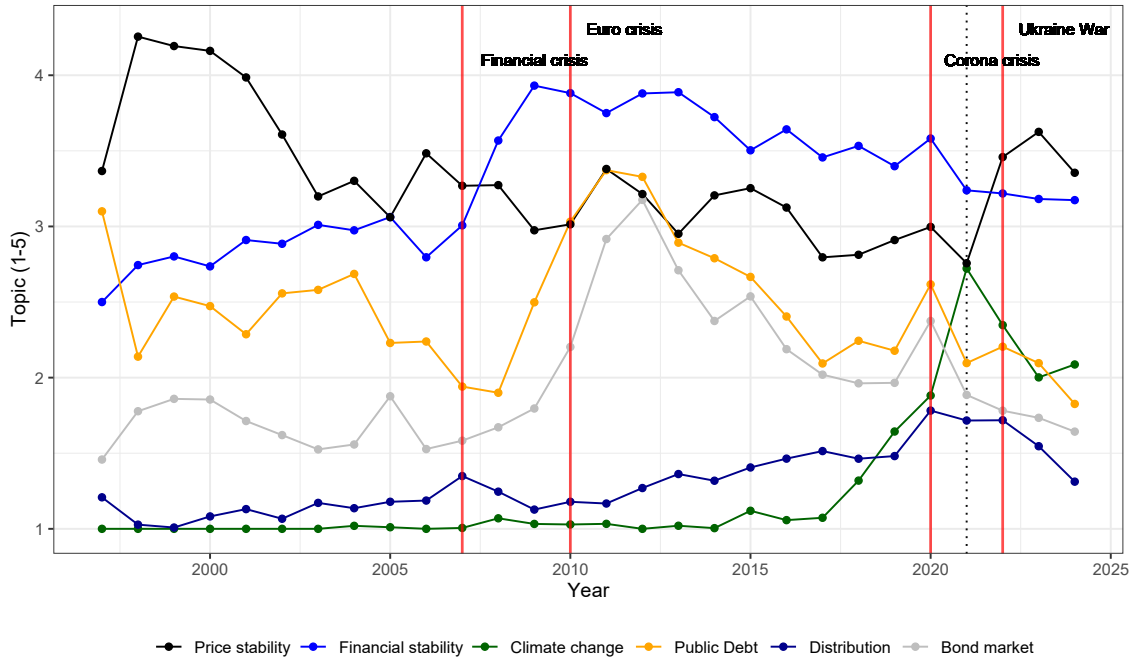
The second main finding is that the discussion on public debt and sovereign bond market stability received a more prominent role during the euro area debt crisis. The upward trend in both categories began already in the financial crisis but only became a major theme after the first signs of the sovereign debt crisis. After the peak of the euro crisis, attention to public debt and sovereign bond markets decreased again.

Third, the topic climate protection entered the public discourse only in the more recent years and then quickly became one of the main themes. In 2014, ECB Council members still almost fully ignored the implications of climate policy for monetary policy. Seven years later,

in 2021, the picture had completely changed and climate issues had become an important part in the debate. The prevalence score for climate protection objective reached its peak in 2021 and sharply declined after that but kept a much higher level compared to the earlier years.

Finally, also the distribution category shows some dynamic over time. In the early years of the euro, terms relating to distribution and inequality rarely occurred. Since 2011, there has been an upward trend in this category. Anecdotal evidence shows that speeches in the first decade that cover the topic do so often in a general sense. For example, Lorenzo Smaghi, former executive board member held a speech in 2007, where he discussed the repercussions of globalization for inequality (Smaghi, 2007). In recent speeches, inequality is often discussed in the light of actual monetary policies, primarily quantitative easing. For instance, in 2021 Jens Weidmann delivered a full speech on inequality, where he discusses the effects of monetary policies on distribution (Weidmann, 2021).

Figure 3: LLM topic prevalence over time

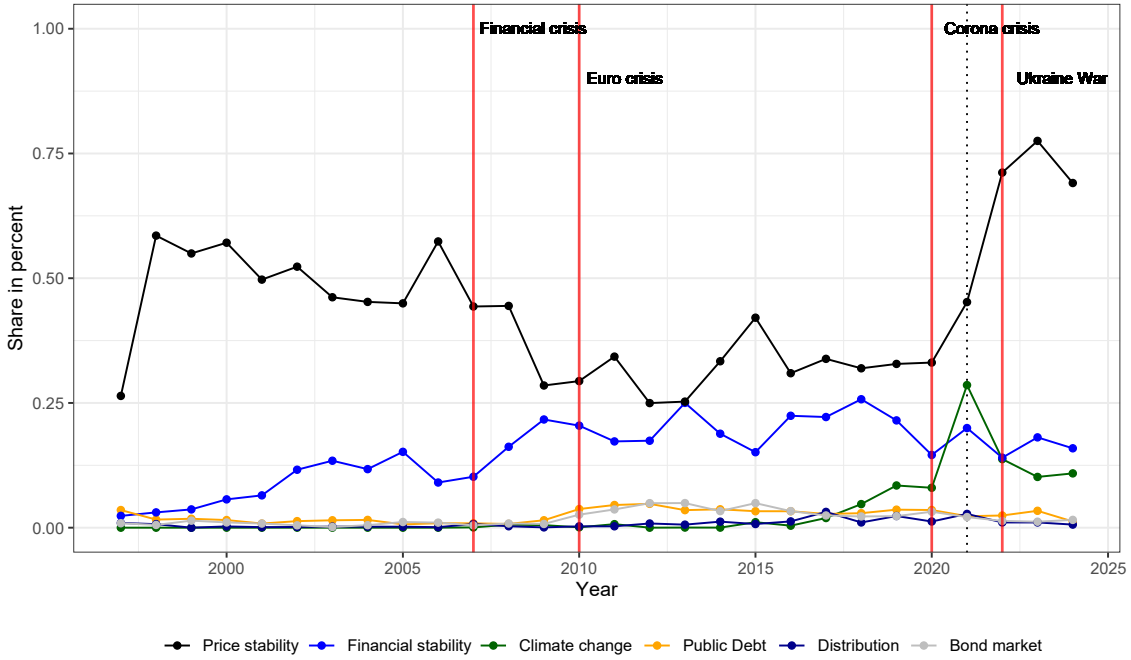


Notes: On the x-axis is the topic prevalence score which ranges from 1 (topic not mentioned) to 5 (speech completely about the topic). The speeches are from the BIS corpus. Each dot shows the mean value in the given year for the given category. The vertical lines indicate the year of relevant crisis in the euro zone.

Figure 4 shows the evolution of the discourse over time according to the prevalence score, this time based on the dictionary analysis. The general trends of the LLM classification are confirmed in the dictionary analysis which strengthens the credibility of the LLM analysis.

The level differences between the categories in the dictionary analysis should be cautiously interpreted because different dictionary categories are likely to capture the objectives to varying degrees.

Figure 4: Dictionary topic prevalence over time



Notes: The x-axis shows the year, while the y-axis reports the annual average share of words in a speech that belongs to the respective dictionary category. The speeches are drawn from the BIS corpus. Each dot represents the mean value for a given category in a given year. The vertical lines indicate the year of relevant crisis in the euro zone.

Topic position

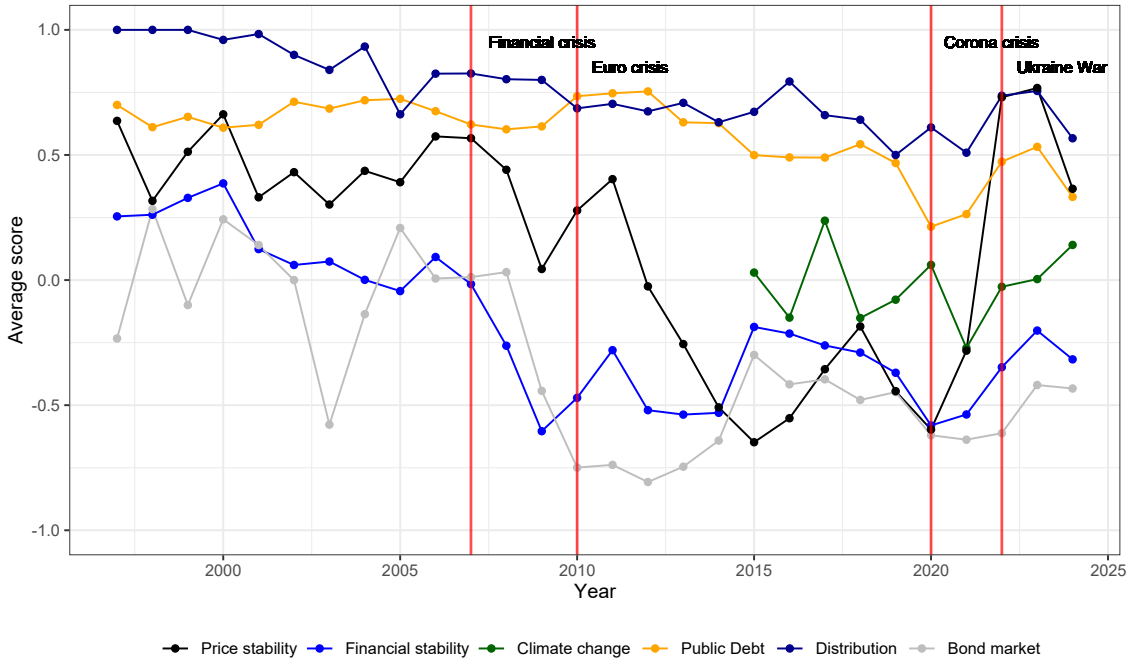
Figure 5 shows the development of the LLM position scores over time for each category with its scale from -1 to 1 as described above. For all objectives, a falling score signals a falling (relative) attention for price stability. The price stability score corresponds to a direct measure of the conventional hawkish-/dovishness dimension, with lower values representing more dovish positions. For the other objectives, the score measures the relative emphasis on price stability relative to the respective objective, with lower values representing a larger weight for the alternative objective relative to price stability.

The time series pattern suggests that the ECB became more dovish after the early years of the euro and developed a more diverse objective function: First, the classical hawkishness indicator declined signalling lower concerns about risks to price stability. Second, other objectives increased in importance relative to the price stability objective. This development was particularly pronounced after 2005. A partial reversal occurred with the onset of high

inflation 2021 when speeches showed a renewed stronger focus on the primary objective of the ECB.

For financial stability the position score shows a lasting change since the financial crisis. The ECB board became more in favour of operations to secure financial stability, whereas in the years before the score was more neutral. The score for climate protection is only available from 2015 onwards due to a lack of coverage before. Since then, the position score fluctuates around a neutral position (score 0.0) with falling support in the most recent, more inflationary years. The emphasis on inequality as a monetary policy objective increased (score decreased) in the years before the pandemic but has returned to a cautious position since then. For the fiscal category public debt the opinion score was relatively stable with a conservative average position emphasizing the priority of price stability. The sovereign bond market opinion score decreased substantially during the euro crisis and again in the years of the pandemic with a trend decline of the average value since the early years of the euro. This shift could represent an increasing permanent attention in the ECB Council for the role of monetary policy for government financing conditions and the relevance of bond markets for monetary policy transmission. This also mirrors the ECB's increasing presence in the markets with the various programmes for government bond purchases established since the first precedent, the Securities Market Programme in 2010.

Figure 5: LLM topic position scores over time



Notes: On the y-axis is the position score which ranges from -1 (strong relative emphasis on alternative objective) to 1 (strong relative emphasis on primary objective). The speeches are from the BIS corpus. Each dot shows the mean value in the given year for the given category. The vertical lines indicate the year of relevant crisis in the euro zone.

4.2 Objectives by country

In this section we compare topic prevalence and positions across countries. The euro member states are a heterogeneous group of countries which vary across many political and economic dimensions. Also, their national central banks have a diverse history and reacted differently to challenges in the pre-euro era (Cukierman et al., 1992). Therefore, it is likely that the euro member states and their representatives bring in different traditions and agendas regarding their preferred monetary policy. In this section, we examine this proposition by analysing the heterogeneity across member states and compared to the ECB board in their speech content.

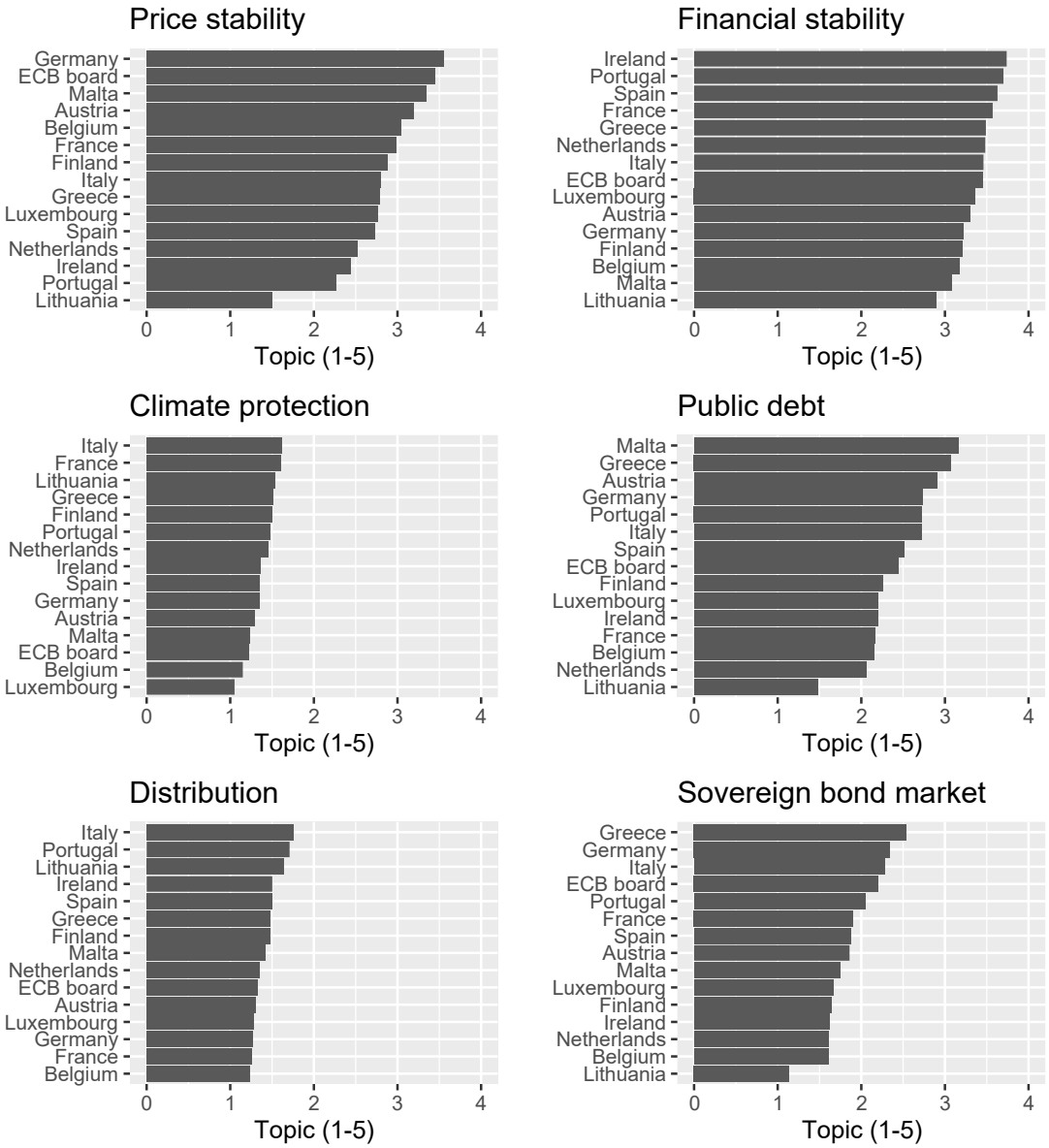
In Figure 6, the average LLM topic prevalence is presented for each member state and the ECB board. The score is averaged over the total observation period (1999-2024) and includes only member states where at least 20 speeches are available in order to avoid unrepresentative results. Nonetheless, the results should be carefully interpreted because the speeches can originate from different time periods under different economic and monetary conditions. In our regression framework below, we tackle this problem by adding varying fixed effects.

In the upper-left corner, the price stability prevalence score is depicted. Germany is leading this category with a score of 3.5 indicting that a large share of speeches covers the objective "price stability" extensively. Germany is followed by the ECB board, Malta and Austria. Most of the Southern European countries have a lower prevalence of price stability in comparison to the Central European countries. Figure A.4 in the appendix also shows the topic position score by country. There, the north-south division, with northern countries having on average a more hawkish position than their southern peers, is clearly visible.

In the upper right corner, the results for the financial stability category are depicted. In general, all countries mention the topic very frequently. At the top are Ireland, Portugal and Spain. According to the topic position score classification (see Appendix Figure A.4), these countries favor a more engaging stance in the realm of financial stability while Germany and Austria are more cautious.

The two fiscal dimensions sovereign bond market stability and public debt are located in the right column of the panel. For the sovereign bond market stability category, Greece and Germany are the most vocal. The topic position score (see Appendix Figure A.3 for the position scores) indicates that Germany adopts a more sceptical tone toward an active ECB role in sovereign bond markets, whereas Greece, on average, expresses a more favourable position. For the public debt category, Germany, Austria and the Southern European countries are most active. The majority of countries adopt a conservative stance toward the objective of public debt, with Germany again having a particularly sceptical position (see Appendix Figure A.3 for the position scores).

Figure 6: LLM topic prevalence by country and ECB board



Notes: The x-axis displays the average topic prevalence score per country which ranges from 1 to 5. Only countries with at least 20 available speeches are included. The title of the subfigures indicates the respective objective.

In the middle-left panel, the results for the climate protection category are shown. Italy, France, Lithuania and Greece are at the top. Lithuania's outcome is slightly misleading because the earliest speech by a Lithuanian governor occurred in 2017, which was already in a period where climate protection entered the stage of central banking. The ECB board, Luxembourg and Belgium are at the bottom of the distribution. These results should be carefully interpreted because climate protection is an emerging topic and for Belgium and Luxembourg only few speeches are included for the relevant observation period. More surprising is the finding that the ECB board scores very low. Considering that climate change was directly addressed with a "climate change action plan" in the ECB 2021 strategy review,

one would expect to see a more active role in the climate domain. Figure A.5 in the Appendix shows, however, that only some ECB board members focus on the implicit objective of climate protection (Figure A.6 shows the respective topic position scores).

Finally, in the bottom left panel the distribution category is displayed. Italy, Portugal, Spain and Greece are more vocal than their Central European peers.

To conclude, there are substantial differences for several of our objectives across countries and the ECB board. For Southern European euro member countries, the price stability category is less prevalent than for the Central European member states. Regarding the debate on sovereign bond market stability, Germany, Greece, Italy and the ECB board are very vocal, although with different positions. Both fiscal topics, public debt and sovereign bond markets, receive a relatively large coverage in Southern Europe.

5 Objectives and national public debt

5.1 Empirical Specification

The variance of positions in ECB Council speeches across time and countries as depicted above can have several explanations. These explanations can be related to the central bank's time-varying political and economic environment, to country characteristics including different national monetary policy views and traditions, and to the individual central banker's characteristics including former inflation experience and individual beliefs and preferences. We focus on the possible connection between national governors' emphasis on certain objectives and the fiscal circumstances in their home country. In particular, we test whether there is an association between the state of public finances in a governor's home country and her coverage of/position on the different possible monetary policy objectives. For this analysis, we restrict the sample to speeches delivered by national governors.

This analysis wants to inform the fiscal dominance debate to which extent the formation of monetary policy positions in a central bank may be overshadowed by considerations related to public debt levels and fiscal sustainability. One possible channel through which fiscal conditions of euro countries may constrain the ECB is that the financing needs of national governments leave an imprint on the monetary policy views of national governors. Hence, we ask to what extent the observable pattern of national public debt and speeches from national governors confirm this channel.

We exploit the panel structure of our data and run panel regressions with varying fixed effects. Depending on the specification, we add year fixed effects to control for the monetary and economic situation in the Eurozone when the speech was held, and country fixed effects to control for time-invariant country characteristics that may impact governors' communication. Our full specification is the following one:

$$y_{ct} = \alpha_c + \gamma_t + \beta \times public\ debt_{ct} + \varepsilon_{ct}$$

y_{ct} is either the topic prevalence or position score for country *c* in year t. α_c is a country fixed effect and γ_t are year fixed effects. The year fixed effect would amongst other factors take account for all the classical determinants (euro area inflation, inflation expectations, output gap) that explain the monetary policy stance of the ECB in the context of an ECB Taylor equation. Furthermore, the year fixed effect controls for general debate trends in the ECB council. $public\ debt_{ct}$ is the public debt level for country *c* in year t. ε_{ct} is the error term. β captures the correlation between public debt and the outcome variable *y* controlled for the euro zone year time trend and country fixed effects. Standard errors are clustered at the country level for all specifications.

The identifying variation in this full model including time and country fixed effects comes from within-country changes in public debt. In an extension, we add individual fixed effect to study the role of national selection of individual governors for changes in national positions.

Our expectations for a fiscal dominance pattern are as follows: an increase in public debt should be associated with an increase in topic prevalence of objectives other than price stability and/or a decrease in price stability as this increases the flexibility of monetary policy decisions. In respect to the position scores, we expect – for a regime of fiscal dominance - that an increase in public debt correlates with a more favourable position toward expansionary monetary policy. For the position scores of the other implicit objectives, we hypothesize that an increase in public debt is associated with a more supportive position for second-order objectives. Emphasizing second-order objectives can be another strategy to downplay the primary objective and to legitimize a larger monetary policy flexibility.

For robustness, we also take the budget balance and the cyclically adjusted budget balance as alternatives for the public debt level.

This two-way fixed effect specification filters out key potential drivers for monetary policy views: First, it controls for the euro area's economic situation including inflation (through the time fixed effects). Second, it controls for time-invariant country characteristics like the monetary policy tradition of a country (through the country fixed effects). Nevertheless, there are obviously other potential time-varying country factors which could affect the validity of the exogeneity assumption. Hence, the results cannot offer a strict proof of causality. However, they do provide evidence whether the association of domestic public debt and revealed objectives in governors' speeches is in line with the expectation for a fiscal dominance regime, where national fiscal interests influence national governors to argue for more monetary policy flexibility. In this sense, our method is of a forensic nature that cannot offer a proof for fiscal dominance but does provide insightful circumstantial evidence.

Furthermore, the availability of speeches varies considerably across countries (see Figure 2), which implies that countries with more extensive speech coverage exert greater influence on the regression estimates.

5.2 Results

Table 2 makes use of our synthetical indicator of the secondary objectives prevalence as the dependent variable. The key result from this first step is that a higher public debt-to-GDP ratio correlates negatively with the secondary objectives index. Specifically, a 10-percentage-point increase in the debt-to-GDP ratio is associated with a 0.015 reduction in the index, corresponding to an 8.1 percent decline relative to its mean value of 0.184. This shows that governors from high-debt countries tend to talk significantly more about secondary objectives relative to the coverage of the price stability objective. This result is robust at the 10 percent significance level when including the full set of time and country fixed effects. It thus points to a within-country pattern: The national governors focus more/less on inflation when their home country's debt-GDP-ratio decreases/increases.

Table 2: Regression results, composite LLM secondary prevalence index

	Secondary Objectives Index	Secondary Objectives Index	Secondary Objectives Index
Debt/GDP	-0.00159**	-0.00098+	-0.00148+
	(0.00047)	(0.00049)	(0.00086)
Year FE	No	Yes	Yes
Country FE	No	No	Yes
Mean dep.	0.184	0.184	0.184
Num.Obs.	2187	2187	2187
R2	0.029	0.132	0.196
Cluster SE	Country	Country	Country

Note: Dependent variable is the secondary objective prevalence index which ranges from -1 (speech entirely about secondary objectives) to 1 (speech only about price stability compared to the selected objectives). In 9 out of the 2,196 speeches from the national governors no classification or a classification error was returned by the algorithm in at least one of the categories. These are dropped from the respective analysis. Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

In the next step, we investigate which particular objectives do drive this overall result for the aggregate objectives indicator. Table 3 shows the results for the individual LLM topic prevalence scores with year fixed effects, while Table 4 additionally includes country fixed effects.

For the secondary objectives climate protection, public debt, sovereign bond markets and distribution, a country's higher debt-to-GDP ratio is associated with more attention for secondary objectives in the speeches. For example, an increase of 100 percentage points in the debt-to-GDP ratio is correlated with a 0.139 increase in the climate protection prevalence score which is a 10-percent increase of the score's mean value. For public debt

and sovereign bond markets, the effects amount to an above 20-percent increase of the mean value. For distribution, the prevalence score increases by 14 percent of the mean if the government-debt-to-GDP ratio increases by 100 percentage points. The distribution category is close to significant at the 10 percent level, while the other mentioned categories are statistically significant at least at the 5 percent level. For price stability, the estimate is negative but not statistically significant. The estimate for financial stability is positive but far from statistical significance.

Table 3: Regression results, LLM topic prevalence, year fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00106	0.00113	0.00139*	0.00466*	0.00204	0.00509*
	(0.00218)	(0.00130)	(0.00066)	(0.00173)	(0.00125)	(0.00181)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	No
Mean dep.	2.88	3.45	1.43	2.48	1.43	1.94
N	2192	2192	2192	2190	2191	2190
R-squarred	0.071	0.101	0.292	0.112	0.080	0.160
Cluster SE	Country	Country	Country	Country	Country	Country

Note: Dependent variable is the topic prevalence score obtained with GPT-4-o mini, which ranges from 1 (objective not mentioned) to 5 (speech entirely on the respective topic). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To go beyond the cross-sectional comparison, we add country fixed effect to exploit the within-country debt-to-GDP variation. Table 4 shows the results. There is a significant and sizable negative correlation between an increase in the public debt-to-GDP ratio in a central banker's home country and the coverage of price stability. A 100 percentage point increase in the debt-to-GDP ratio is correlated with a 22 percent decrease in the mean prevalence score for price stability. Conversely, an increasing country debt level is linked to a higher prevalence of climate protection and distribution in speeches but only the effect for distribution is significant. Thus, the more demanding two-way fixed effect test does not confirm a strong link between debt and prevalence for each single objective in isolation with the exception of distribution (Table 4), but it does confirm the link for the bunch of secondary objectives together (Table 2). Central bankers start to talk more about other objectives when their home debt increases, but – apart from the distribution topic – this cannot be pinned down to single objectives.

Table 4: Regression results, LLM topic prevalence scores, –two-way fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00665*	-0.00361	0.00012	-0.00108	0.00320*	-0.00224
	(0.00298)	(0.00330)	(0.00205)	(0.00274)	(0.00128)	(0.00187)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep.	2.88	3.45	1.43	2.48	1.43	1.94
N	2192	2192	2192	2190	2191	2190
R-squared	0.141	0.148	0.306	0.155	0.109	0.202
Cluster SE	Country	Country	Country	Country	Country	Country

Note: Dependent variable is the topic prevalence score obtaines with GPT-4-o mini, which ranges from 1 (objective not mentioned) to 5 (speech entirely on the respective topic). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

The topic prevalence analysis only covers to which extent national governors talk about the respective objectives. It does not provide any information on their position toward these objectives. To analyse the contents, we run the same regressions as before, now with the topic position score as the outcome variable. The year-fixed effect specifications (Table 5) show that, for all objectives, a higher debt level is associated with a relatively lower emphasis on price stability relative to the alternative objective. Among the regressions, only those for price stability and the sovereign bond market are estimated with statistical precision. An increase of the debt-to-GDP ratio by 100 percentage points is associated with a 0.4-decrease of the price stability score (range -1.0 to 1.0), which is a substantial magnitude considering that the mean value is 0.15. A lower price stability score expresses a more dovish position. Hence, an increase of the debt-to-GDP ratio, controlling for the economic and monetary condition in the currency union through year fixed effects, is correlated with more dovish statements. Equally, speeches from countries with a higher debt-to-GDP ratio signal a significantly higher support for the ECB to engage in the sovereign bond market.

The number of speeches considered per category varies substantially because the topic position score is defined only for speeches that mention the respective objective. The lowest numbers are for speeches referring to climate and distribution, possibly explaining the low statistical precision of the estimates.

Table 5: Regression results, LLM topic position, year fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00414*	-0.00308	-0.00106	-0.00077	-0.00001	-0.00205***
	(0.00178)	(0.00215)	(0.00167)	(0.00088)	(0.00107)	(0.00034)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	No
Mean dep.	0.152	-0.340	-0.042	0.585	0.600	-0.574
N	1400	1970	372	1204	504	715
R-squared	0.409	0.095	0.083	0.096	0.081	0.231
Cluster SE	Country	Country	Country	Country	Country	Country

Note: The dependent variable is the GPT topic position score calculated by GPT-5 mini, which ranges from -1.0 (low relative attention for prices stability, dovish) to +1,0 (high relative attention for price stability, hawkish). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 6 shows estimation results for the two-way fixed effect specification. The negative signs of the public-debt coefficients for all objectives prove to be robust. Also, the significant result for the price stability category survives this test. The effect is similar in size which shows that there is a substantial within-country link between a country's changing debt-to-GDP ratio and a changing emphasis on price stability of its representatives in the ECB Council. This finding corresponds to the fiscal dominance hypothesis that national presidents become more dovish when their home country's public debt situation deteriorates. For the other categories we find the significant results that higher debt levels (within-country) go hand in hand with a larger emphasis on both the climate protection and the distribution objective relative to price stability.

Table 6: Regression results, LLM topic position, two-way fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00487**	-0.00272	-0.01398*	-0.00071	-0.00595*	-0.00008
	(0.00158)	(0.00224)	(0.00484)	(0.00130)	(0.00226)	(0.00209)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep.	0.152	-0.340	-0.042	0.585	0.600	-0.574
Num.Obs.	1400	1970	372	1204	504	715
R2	0.536	0.320	0.270	0.193	0.234	0.258
Cluster SE	Country	Country	Country	Country	Country	Country

Note: The dependent variable is the GPT topic position score calculated by GPT-5 mini, which ranges from -1.0 (low relative attention for prices stability, dovish) to +1,0 (high relative attention for price stability, hawkish). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

5.3 Robustness checks

5.3.1 Alternative fiscal indicators

As a robustness check, we ran all specifications from above also with two alternative fiscal indicators, the general government budget balance and the cyclically adjusted budget balance as outcome variables. Table 7 shows the results for the topic and the score analysis, the first two columns repeat the results for the debt-to-GDP ratio from above as a reference. In general, we would expect to find the opposite sign for the budget balance in comparison to the public debt level because an increase of the government balance stands for an improving fiscal situation. In the prevalence analysis, all the government balance estimates with year fixed effects in column 3 (non-adjusted) and 5 (cyclically adjusted) have the same direction and for the public debt-to-GDP ratio estimates the sign, as expected, is reversed. Like for the stock of public debt also for the flow variable budget balance, national governors talk less about price stability and more about the secondary objectives if they come from countries with an unfavourable fiscal situation. Contrary to some significant findings for the debt level, the coefficients for the budget balance are not estimated with statistical precision. The clearer result for the debt level compared to the government balance is not too surprising. The latter rather provides a snapshot on a country's current fiscal situation whereas the former is indicative of the remaining fiscal space. The two-way fixed effect estimates do not provide a clear picture with switching signs and a lack of statistical significance.

Table 7: Regression results, alternative fiscal indicators, LLM topic prevalence

	Independent variables					
Objective	Debt GDP ratio		Budget balance		Cyclically adjusted budget balance	
Price stability	-0.0011	-0.0066*	0.0082	-0.0077	0.0118	-0.0069
	(0.0022)	(0.0030)	(0.0214)	(0.0140)	(0.0240)	(0.0171)
Financial stability	0.0011	-0.0036	-0.0077	0.0192	-0.0092	0.0249
	(0.0013)	(0.0033)	(0.0103)	(0.0151)	(0.0130)	(0.0156)
Climate protection	0.0014*	0.0001	-0.0073	-0.0063	-0.0068	-0.0037
	(0.0007)	(0.0020)	(0.0079)	(0.0049)	(0.0095)	(0.0061)
Public debt	0.0047*	-0.0011	-0.0214	-0.0215+	-0.0219	-0.0285*
	(0.0017)	(0.0027)	(0.0167)	(0.0121)	(0.0165)	(0.0112)
Distribution	0.0020	0.0032*	-0.0042	0.0002	-0.0031	0.0010
	(0.0013)	(0.0013)	(0.0094)	(0.0059)	(0.0120)	(0.0076)
Sovereign bond market	0.0051*	-0.0022	-0.0046	0.0003	-0.0047	0.0011
	(0.0018)	(0.0019)	(0.0160)	(0.0125)	(0.0158)	(0.0135)
Fixed effects	Year	Two-way	Year	Two-way	Year	Two-way

Note: Dependent variable is the prevalence score obtained with GPT-4-o mini, which ranges from 1 (objective not mentioned) to 5 (speech entirely on the respective topic). Cells include regression coeffcients. Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

For the position score (Table 8), the budget-balance variable shows a significantly positive relationship (in the two-way regression) with both the climate-protection and the public

debt topic score, indicating that larger fiscal surpluses are associated with a stronger focus on the primary objective. The same holds for the traditional hawkishness indicator (price stability position score). These findings are largely consistent with the results obtained from the debt-to-GDP ratio analysis. The results for the cyclically adjusted budget balance variable are all insignificant, except for the public debt objective.

Table 8: Regression results, alternative fiscal indicators, LLM position score

	Independent variables					
Objective/	Debt GDP ratio		Budget balance		Cyclically adjusted budget balance	
Price stability	-0.0041*	-0.0049**	0.0460+	0.0249*	0.0366	0.0128
	(0.0018)	(0.0016)	(0.0238)	(0.0117)	(0.0233)	(0.0145)
Financial stability	-0.0031	-0.0027	0.0303	-0.0017	0.0290	-0.0092
	(0.0022)	(0.0022)	(0.0238)	(0.0065)	(0.0224)	(0.0086)
Climate protection	-0.0011	-0.0140*	0.0167	0.0451+	0.0128	-0.0016
	(0.0017)	(0.0048)	(0.0250)	(0.0229)	(0.0217)	(0.0310)
Public debt	-0.0008	-0.0007	0.0119+	0.0038+	0.0128+	0.0036
	(0.0009)	(0.0013)	(0.0066)	(0.0022)	(0.0069)	(0.0044)
Distribution	-0.0000	-0.0060*	0.0070	0.0087	0.0112	0.0123
	(0.0011)	(0.0023)	(0.0072)	(0.0064)	(0.0075)	(0.0085)
Sovereign bond market	-0.0021***	-0.0001	0.0022	-0.0036	0.0009	-0.0042
	(0.0003)	(0.0021)	(0.0042)	(0.0066)	(0.0063)	(0.0075)
Fixed Effects	Year	Two-way	Year	Two-way	Year	Two-way

Dependent variable: positioning GPT score. Cells include regression coeffcients. Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To summarize, for the topic analysis with the budget balance as the indicator for a country's fiscal health the effects mostly point in the same direction as with the debt level. However, the results for the topic prevalence are rarely statistically significant while the position score results are more reliable. The somewhat stronger results for the debt level are plausible if one regard the stock of debt as the more informative indicator for a country's structural fiscal situation than the government balance.

5.3.2 Dictionary results

Table A.4 and Table A.5 in the Appendix report the results of the dictionary-based validation checks for topic prevalence. Table A.4 presents the year fixed-effects regressions. The coefficient for public debt is positive and statistically significant, indicating that a 100-percentage-point increase in the debt-to-GDP ratio is associated with a 0.019 increase in the share of public-debt-related dictionary terms in a speech. The coefficient for climate protection is also positive, although not statistically significant. For the distribution and sovereign bond market categories, the estimated effects are small and not significant. For the price stability objective, we observe a negative correlation with the debt-to-GDP ratio, which is close to significance at the 10 percent level (p = 0.105). Table A.5 reports the results using two-way fixed effects. Here, the coefficient for price stability is again negative but not

statistically significant. The estimated effects for the implicit objectives are small and imprecisely estimated overall.

To summarize, several of the dictionary-based results point in the same direction as the LLM analysis: speeches tend to cover secondary objectives more and the primary objective less with rising debt levels in the speaker's home country. However, these results lack the statistical precision required for reliable inference. This may be due to the dictionary's limited ability to precisely capture the underlying concepts.

6 Mechanism

6.1 Political selection

So far we have studied to which extent a country's governors express views that correspond to the home country's fiscal situation but have not looked deeper into the mechanism. There are two main channels through which a country's central bankers' positions may shift in response to the level of public debt. The first is governor selection. Even in the euro setting with central bank independence guaranteed by EU primary law, the governments select and appoint the governors of their national central bank. Facing higher debt, a government might be more likely to appoint a central banker who supports a more accommodative monetary stance.

The second channel is that an incumbent governor may adjust his views and communication over time. For example, the same governor might adopt a more hawkish stance as the fiscal situation in his home country improves.

In this extension, we distinguish between both mechanisms and replace the country fixed effect with a governor individual fixed effect. Hence, we only exploit individual variation for the same central banker which basically tests if his position has changed over time. If significant estimates for the country fixed effects do no longer show up with individual fixed effects the conclusion would be that governor selection drives our results.

For the objectives' prevalence score we find smaller and no longer significant effects (Table 9 compared to Table 4) which points in the direction that a part of the previous found effects is driven by governor turnover. For the position score, the results for the hawkish-dovish dimension remain robust and the point estimator barely changes (Table 10 compared to Table 6). The same holds true for the climate protection and distribution category. These results provide some evidence in the direction that individual central bankers' positions toward the objectives are directly or indirectly affected and that, hence, national governor positions do change without changes in the leadership of the central banks. On the other side, the content composition of speeches seems to be partly driven by a selection of the national central bank presidents.

Table 9: Regression results, LLM topic prevalence - individual fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00123	0.00042	0.00383	0.00367	-0.00353	0.00366
	(0.00280)	(0.00205)	(0.00350)	(0.00428)	(0.00306)	(0.00363)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	No
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep.	2.88	3.45	1.43	2.48	1.43	1.94
N	2192	2192	2192	2190	2191	2190
R-squared	0.195	0.194	0.320	0.207	0.138	0.239
Cluster SE	Country	Country	Country	Country	Country	Country

Note: Dependent variable is the topic prevalence score obtained by GPT-4-o mini, which ranges from 1 (objective not mentioned) to 5 (speech entirely on the respective topic). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 10: Regression results, LLM topic position - individual fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00497+	-0.00138	-0.01279*	0.00024	-0.00716*	0.00247
	(0.00274)	(0.00186)	(0.00527)	(0.00119)	(0.00325)	(0.00242)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	No
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep.	0.152	-0.340	-0.042	0.585	0.600	-0.574
Num.Obs.	1400	1970	372	1204	504	715
R2	0.595	0.396	0.281	0.261	0.289	0.345
Cluster SE	Country	Country	Country	Country	Country	Country

Note: Dependent variable is the topic position score calculated by GPT-5 mini, which ranges from -1.0 (low relative attention for price stability, dovish) to 1.0 (high relative attention for price stability, hawkish). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

6.2 Nonlinearities

So far, we only examined a linear correlation between the central bankers' positions and public debt. But already the robustness checks with the government balance pointed in a direction of potential nonlinearities. It might be that additional debt at a critical public debt level has a stronger effect on monetary policy than the same debt, when the initial level is low. To test this, we add the squared debt-to-GDP ratio to our main model. The results are presented in Table 11. For the composite prevalence score (column 1), the estimated coefficients for debt-to-GDP and debt-to-GDP squared point into the hypothesis supporting direction but fall short of conventional levels of statistical significance.

For the price stability position score (column 2), we find evidence for nonlinearities. An increasing level of public debt leads to a stronger dovish effect when the initial public debt level is higher. This finding supports the hypothesis that, at unsustainable levels of public

debt, the political pressure on central bankers becomes stronger. For the public debt position score (column 5), we find an interesting non-linearity. With increasing public debt, central bankers have a more conservative stance on the topic but at very high levels of public debt the position becomes more dovish.

Table 11: Nonlinearities, LLM topic position - two-way fixed effects

	Secondary Objectives Index	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00011	-0.00074	0.00067	-0.01617+	0.00361+	-0.00665	0.00201
	(0.00297)	(0.00244)	(0.00479)	(0.00906)	(0.00189)	(0.00561)	(0.00505)
Debt/GDP squared	-0.00001	-0.00002+	-0.00002	0.00001	-0.00002**	0.00000	-0.00001
	(0.00001)	(0.00001)	(0.00003)	(0.00003)	(0.00001)	(0.00002)	(0.00002)
N	2187	1400	1970	372	1204	504	715
R-squared	0.196	0.537	0.321	0.270	0.198	0.234	0.258
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Country	Country	Country	Country	Country	Country	Country

Note: Dependent variable in column 1 is the secondary objective prevalence index which ranges from -1 (speech entirely about secondary objectives) to 1 (speech only about price stability compared to the selected objectives) classified by the GPT-4o mini model. For column 2 to 7 the outcome variable is the respective topic position score calculated by GPT-5 mini, which ranges from -1.0 (low relative attention for price stability, dovish) to 1.0 (high relative attention for price stability, hawkish). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

7 Conclusion

In this paper, we analyse 4,680 speeches given by ECB Council members since the introduction of the euro, focusing on measuring attention given to different monetary policy objectives. Our findings suggest that the discourse is a dynamic process that has become more diverse in recent years, covering an increasingly broad spectrum of objectives. The financial and euro crises have left their mark on central bank discourse, with financial stability, public debt, and sovereign bond market stability receiving more attention subsequently. Until 2015, climate considerations were almost entirely overlooked by Council members. However, since 2017, this issue has rapidly gained attention and has become an important topic in the discourse. Distributional concerns have also received more attention since the financial crisis. Additionally, the recent rise in inflation has influenced external communication, with ECB Council members becoming much more hawkish on average and showing renewed concern for the primary objective of price stability.

Furthermore, our analysis shows that there is considerable heterogeneity among member states and individual central bankers regarding their public discussion of objectives. We are

primarily interested in searching for a systematic relationship between a country's fiscal situation and the objective function of its national governor, as revealed in their speeches. In order to inform the fiscal dominance debate, we examine whether a less favourable domestic fiscal situation is indicative of a governor who places greater emphasis on alternative objectives than on the primary objective.

We ran several one-way and two-way panel regressions, examining both the coverage and support of objectives (i.e. topic prevalence and topic position). The results confirm that higher debt is firstly associated with greater attention to secondary objectives, even in the presence of time and country fixed effects. Secondly, higher debt is associated with stronger support for secondary objectives relative to primary objectives. This effect is particularly pronounced for the positions on climate protection and distribution, with significances in the two-way specification establishing a correlation within countries: an increase in a euro country's national debt goes hand in hand with its national governor shifting attention away from price stability towards climate and distribution policy.

We conduct a number of validation tests to verify the reliability of the LLM-generated scores. For a sub-sample, we compare LLM scores with manually coded scores. We also assess the trajectories of our prevalence and position scores against contemporaneous macro developments and provide a traditional, dictionary-based prevalence score for comparison. These validation exercises bolster our confidence in the information content of the LLM scores.

Overall, the results suggest that elevated public debt in a governor's home country is associated with a shift in policy priorities, moving away from price stability towards secondary objectives. Selection in central bank appointments contributes to this relationship, but does not fully explain it.

As with any panel-based identification strategy, our results have natural limitations. Although we control for year and country fixed effects, there are still many potential confounders from time-varying national factors. Therefore, the results show correlations and should be interpreted with the usual caution. However, our finding that growing domestic public debt levels are associated with national governors who exhibit in their communication a more diverse monetary policy objective function and a lower attention for price stability is an important observation for the fiscal dominance debate.

References

Apel, M., Blix Grimaldi, M., Hull, I., 2022. How much information do monetary policy committees disclose? Evidence from the FOMC's minutes and transcripts. Journal of Money, Credit and Banking 54, 1459–1490.

Apel, M., Grimaldi, M.B., 2014. How informative are central bank minutes? Review of Economics 65, 53–76.

Bank for International Settlements, 2024. Central Bankers' speeches.

Baumgärtner, M., Zahner, J., 2021. Whatever it Takes to Understand a Central Banker – Embedding their Words Using Neural Networks.

Bennani, H., Neuenkirch, M., 2017. The (home) bias of European central bankers: new evidence based on speeches. Applied Economics 49, 1114–1131.

Bjerkander, L., Glas, A., 2024. Talking in a language that everyone can understand? Clarity of speeches by the ECB Executive Board. Journal of International Money and Finance 149, 103200.

Boffa, C., Treeck, J., 2025. ECB's Lagarde set to face grilling on French troubles. POLITICO.

Bohl, M.T., Kanelis, D., Siklos, P.L., 2023. Central bank mandates: How differences can influence the content and tone of central bank communication. Journal of International Money and Finance 130, 102752.

Bordo, M.D., Levy, M.D., 2021. Do enlarged fiscal deficits cause inflation? The historical record. Economic Affairs 41, 59–83.

Chappell Jr, H.W., Havrilesky, T.M., McGregor, R.R., 1997. Monetary policy preferences of individual FOMC members: a content analysis of the memoranda of discussion. Review of Economics and Statistics 79, 454–460.

Chappell Jr, H.W., McGregor, R.R., Vermilyea, T.A., 2008. Regional economic conditions and monetary policy. European Journal of Political Economy 24, 283–293.

Christine Lagarde, 2021. Keynote speech by Christine Lagarde, President of the ECB, at the ILF conference on Green Banking and Green Central Banking.

Cochrane, J., 2023. The fiscal theory of the price level.

Cukierman, A., Web, S.B., Neyapti, B., 1992. Measuring the Independence of Central Banks and Its Effect on Policy Outcomes. World Bank Econ Rev 6, 353–398.

De Haan, J., Eijffinger, S.C.W., 2016. The Politics of Central Bank Independence. De Nederlandsche Bank Working Paper No. 539.

Dossche, M., Slačálek, J., Wolswijk, G., 2021. Monetary Policy and Inequality.

ECB, 2014. ECB assumes responsibility for euro area banking supervision.

European Central Bank, 2022. The Transmission Protection Instrument.

Fadda, P., Hanifi, R., Istrefi, K., Penalver, A., 2025. Central bank communication of uncertainty. Journal of International Money and Finance 157, 103385.

Gardner, B., Scotti, C., Vega, C., 2022. Words speak as loudly as actions: Central bank communication and the response of equity prices to macroeconomic announcements. Journal of Econometrics 231, 387–409.

Gentzkow, M., Kelly, B., Taddy, M., 2019. Text as data. Journal of Economic Literature 57, 535–574.

Hansen, S., McMahon, M., Prat, A., 2018. Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach*. The Quarterly Journal of Economics 133, 801–870.

Hassan, T.A., Hollander, S., Kalyani, A., Van Lent, L., Schwedeler, M., Tahoun, A., 2025. Text as Data in Economic Analysis. Journal of Economic Perspectives 39, 193–220.

Havlik, A., Heinemann, F., 2021. Sliding Down the Slippery Slope? Trends in the Rules and Country Allocations of the Eurosystem's PSPP and PEPP. CCM 54, 173–197.

Heinemann, F., Kemper, J., 2021. The ECB under the threat of fiscal dominance–The Individual Central Banker Dimension. The Economists' Voice 18, 5–30.

Hinterlang, N., Hollmayr, J., 2020. Classification of monetary and fiscal dominance regimes using machine learning techniques, Discussion paper. Deutsche Bundesbank, Frankfurt am Main.

Howarth, D., Quaglia, L., 2014. The Steep Road to European Banking Union: Constructing the Single Resolution Mechanism. J Common Mark Stud 52, 125–140.

Kokoszczyński, R., Mackiewicz-Łyziak, J., 2024. Making monetary policy in Poland: Are Polish hawks and doves different? European Journal of Political Economy 81, 102481.

Kronberger Kreis, 2021. The monetary policy strategy of the European Central Bank: Review and recommendations.

Malmendier, U., Nagel, S., Yan, Z., 2021. The making of hawks and doves. Journal of Monetary Economics 117, 19–42.

Pennington, J., Socher, R., Manning, C., 2014. GloVe: Global Vectors for Word Representation.

Picault, M., Renault, T., 2017. Words are not all created equal: A new measure of ECB communication. Journal of International Money and Finance 79, 136–156.

Sargent, T.J., Wallace, N., 1981. Some unpleasant monetarist arithmetic.

Schnabel, I., 2021. From Green Neglect to Green Dominance?

Schnabel, I., 2020. The shadow of fiscal dominance: Misconceptions, perceptions and perspectives.

Shapiro, A.H., Wilson, D.J., 2022. Taking the fed at its word: A new approach to estimating central bank objectives using text analysis. The Review of Economic Studies 89, 2768–2805.

Smaghi, L., 2007. Dinner speech: Globalisation and public perceptions.

Weidmann, J., 2021. Will Covid-19 increase economic inequality.

Woodhouse, D., Charlesworth, A., 2023. Can ChatGPT Predict Future Interest Rate Decisions? SSRN Working Paper.

World Bank, 2022. Inflation, consumer prices (annual %) - Euro area.

Wyplosz, C., 2019. Limits to the independence of the ECB. Rev World Econ 155, 35–41.

Appendix

A.1 Scoring methods

A.1.1 Development of a central bank specific dictionary

A dictionary approach requires a high quality and context specific dictionary which correctly measures the intended concept. To ensure valid results, we develop the dictionary in a multistage semi-supervised procedure combining machine learning methods and qualitative expertise.

The dictionary development is initiated by deriving relevant topics from the academic and public discourse on the ECB's objective function. Then, we identify key terms which capture the respective topic and take them as a starting point. For example, in the case of the "climate protection" category "climate change" and "global warming" were among the initial terms. A common critique of dictionary approaches is their subjectivity in the term selection. When a dictionary is just based on expert knowledge, the potential bias of the expert can be translated into the dictionary. Furthermore, important terms can just be forgotten in the creation process. To tackle these coverage issues, we finetune a deep neural network on the speeches data set which returns sophisticated word embeddings (Pennington et al., 2014).

Word embeddings are numerical vectors which contain meaningful information on the respective word (see e.g. Gentzkow et al., 2019; Pennington et al., 2014). For example, the word "inflation" is likely to co-occur in a context where "prices" are also mentioned. Based on the assumption that words with similar meaning are closer to each other, these models can calculate n-dimensional vectors for every word in their training set. Intuitively these vectors can be thought of as location parameters with very useful characteristics. Research has shown that words which have a similar meaning are close to each other in the vector space (Pennington et al., 2014). For instance, the words "inflation" and "prices" are in a close proximity which can simply be calculated by cosine similarities (Pennington et al., 2014). Word embeddings are generated through an iterative process on a large corpus of text. One can either use pre-trained word embeddings or train them on a specific form of text. As central bank speeches are very different from other text sources where these models are usually trained on, we fine tune our own embeddings based on the complete set of speeches in our data set. Baumgärtner and Zahner (2021) have shown that Glove word embeddings return meaningful results when they are trained on central bank speeches. Therefore, we follow their approach and also train GloVe embeddings (Pennington et al., 2014).

With these word embeddings, we can check which words have a similar meaning to our initial dictionary terms. Then, we qualitatively examine if the words with a similar vector fit into the dictionary. This has the great advantage, that words which were overseen in the initial phase can be detected technically. For example, we start with the word "climate" for the climate dictionary and the word "risk" has a similar vector to "climate", hence, it is likely that "climate" and "risk" often appear together. Hence, the expression "climate risk" is a

potential candidate for the “climate change” category. In the next step, we examine if the new found expression appears frequently in our data set. If the term is just used once in the whole corpus, which is often the case for two word combinations, the inclusion makes no real difference for the results of the analysis. On the other side, terms which are frequently used have a substantial weight on the final results. For these reasons, we want to guarantee that these expressions really represent the underlying topic. Therefore, we do a keyword in context analysis, where we define a window of five words before and after the expression. Then, we first qualitatively check if the context fits to the topic. Afterwards, the most frequent terms in this window across all occurrences are computed. If the most common words are still related to the topic, the new term is incorporated into the dictionary. Furthermore, through this approach more related terms are detected and after testing included in the dictionary. Following this data-driven approach the list of new terms is examined by a domain expert and refined if necessary. Finally, the dictionary is validated by human annotation to test if the automated dictionary topic classification and a human assessment agree.

Table A.1 shows the dictionary. It consists of six categories and only the first ten terms of each category are shown here. In the first column, the objective “price stability” is depicted. The first term which describes this topic is “inflation*”. The asterisk (“*”) is an escape character which implies that whatever follows the previously defined series of characters is counted until a space occurs. So “inflation*” includes also the term “inflationary”. Further terms in this category are “deflation*” and “price stability”. For price stability, the dictionary item “inflation” is used 25,503 times in all speeches and “price stability” occurs 11,692 times. These both terms are the main driver of this category.

Table A.1: Dictionary

Price stability	Financial /banking stability	Public debt	Sovereign bond market stability	Climate protection	Distribution
inflation*	financial stability	national debt*	bond market stress	climate change	income distribution
deflation*	banking stability	public debt*	sovereign yield spread*	global warming	wealth distribution
price stability	banking union	government debt*	market segmentation	green bonds	profit distribution
primary mandate	banking crisis	government's debt	market fragmentation	green finance*	inequality
price developments	deposit protection	public deficit*	bond market stress	green bond*	re-distribution
consumer price*	capital buffer*	government deficit*	debt market stress	green deal	redistribution
energy price*	bank equity	debt sustainability	sovereign risk	green investment*	equality of opportunity
commodity price*	credit risk*	Fiscal space	national bond*	green economy	inclusive growth
hicp	systemic bank*		government bond*	green transition	
price*	systemic risk*			green asset*	

Notes: The table shows dictionary terms for each objective. The asterisk indicates that all whatever follows the previously defined series of characters is counted until a space occurs. The table only shows the top ten term of each category.

A.1.2 LLMs model selection

Among the available LLMs during the analysis phase of the project OpenAI's GPT models were among the top performing models across a wide range of different tasks. Furthermore, the model's high quality performance is widely known through its easily accessible ChatGPT platform. For these reasons, we opt for a GPT model for the classification task. To decide for a concrete model among the GPT models we considered classification quality, run time and costs. We qualitatively examined the classification quality by analysing individual test cases with varying models and checked if the general results are in line with known macroeconomic trends (e.g. see Figure 3).

Furthermore, we generated 154 three sentence paragraphs with GPT-4o about price stability. 50 paragraphs should be dovish, 50 hawkish and 50 neutral. Then, we manually classified the paragraphs and tested several GPT models on them (GPT-5, GPT-4o mini, GPT-5 mini). With the artificial generation, we tested if GPT models have some understanding of central bank lingua and key concepts like hawks and doves.[4] Furthermore, we used the dataset to test how different GPT models perform in stylized and easily comparable cases (the validation results on the real sample are in section 3.4). According to our judgment, GPT-4o generated mostly realistic central bank statements with correctly defined positions along the dovish-hawkishness dimension. Table A.2 shows the result of the pretesting exercise for the category price stability. For the topic classification, the root mean squared error (RMSE) is the lowest for GPT-4o mini. Furthermore, this model is the cheapest among the high performing models (in August 2025) and is very fast. For the position score, the GPT-5 model has the lowest RMSE but it has a high number of missing classification (which cannot be considered in the performance statistics). Qualitatively, we find that GPT-5 struggles to classify the dovish sentences correctly. GPT-5 mini has only a slightly higher RMSE but much less missing classification. GPT-4o mini performs poorly with correctly assessing dovish sentences. Hence we choose GPT-5 mini as our model for the position score classification.

[4] The generate sentences come from a different model GPT-4o than the models with use to pre-test them GPT-4o mini, GPT-5 and GPT-5 mini.

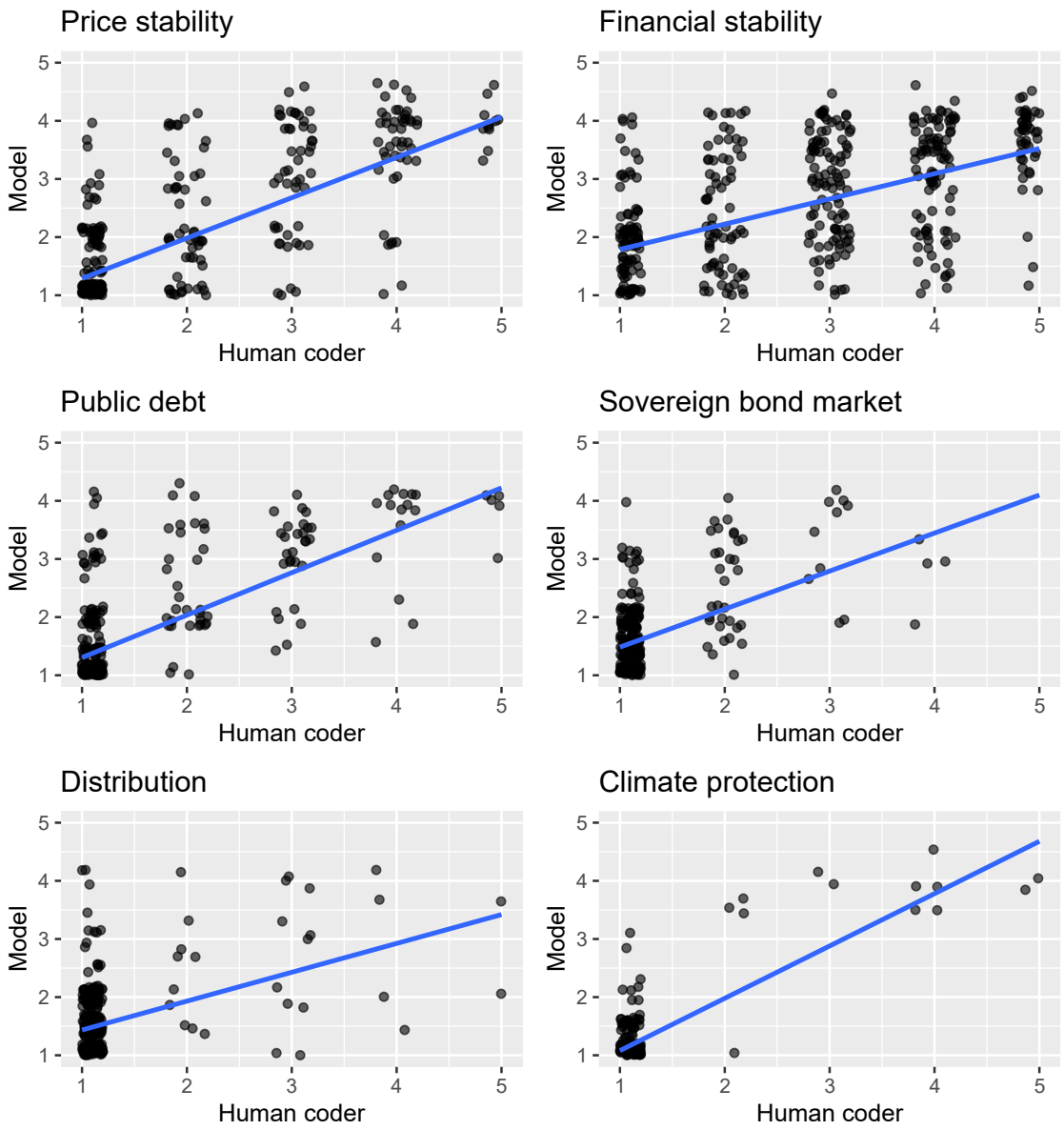
Table A.2: Pre-test model selection - generated data

Model	Accuracy	F1 Macro	NA Count	Precision Macro	Recall Macro	RMSE
Score GPT-5o	0.925	0.892	48	0.915	0.876	0.257
Score GPT-4o Mini	0.813	0.592	20	0.551	0.640	0.475
Score GPT-5o Mini	0.875	0.847	34	0.836	0.871	0.272
Topic GPT-5o	0.325	0.242	0	0.331	0.373	1.349
Topic GPT-4o Mini	0.357	0.274	0	0.344	0.311	0.983
Topic GPT-5o Mini	0.409	0.330	0	0.424	0.408	1.219

Notes: The data consists of 154 paragraphs with three sentences each on the objective of price stability. The same prompt was sent to all models. The continuous results were rounded to integer values to calculate classification statistics (not applied for root mean squared error). The NA count column indicates the number of statements which could not be classified.

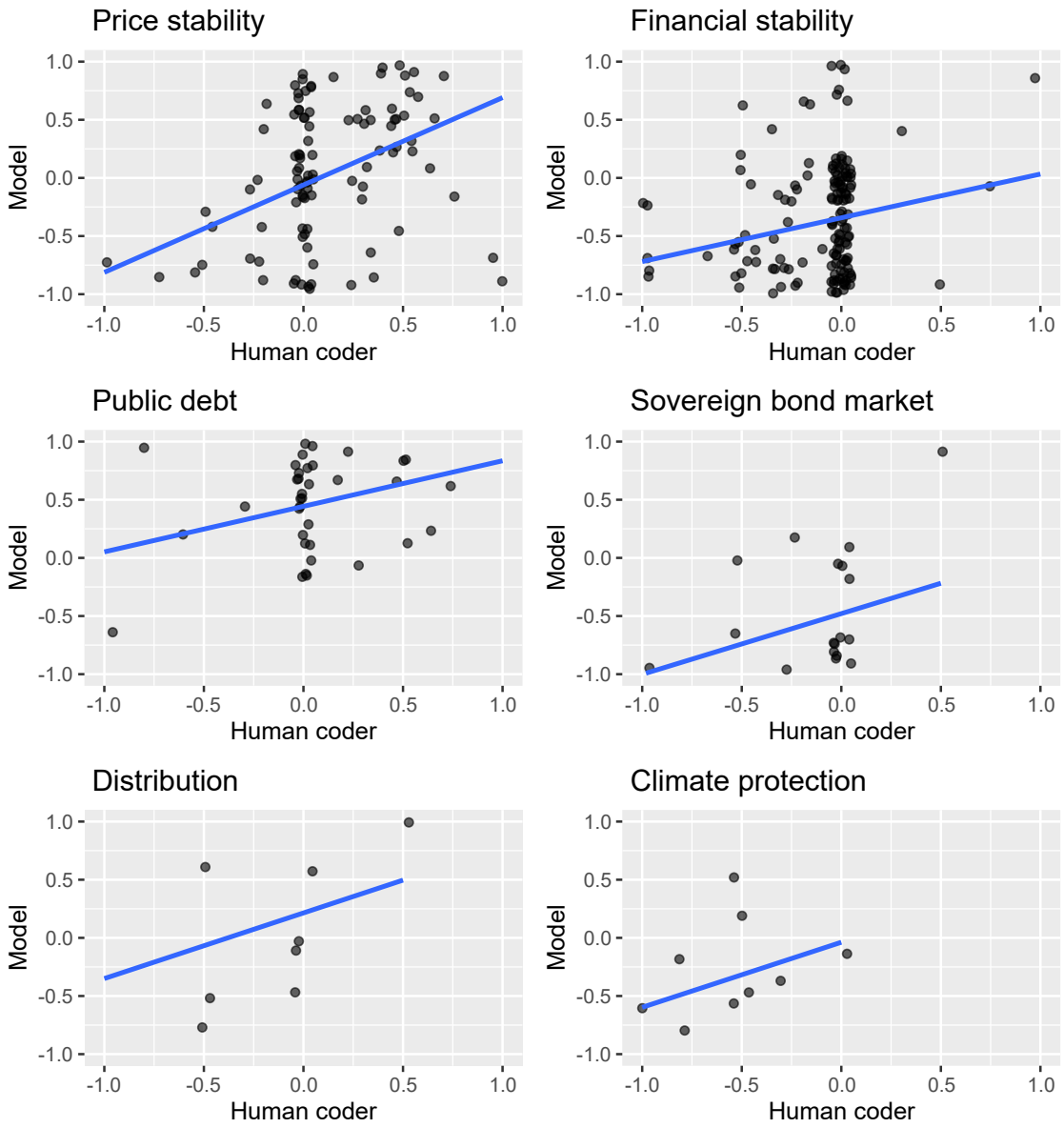
A.1.3 LLM validation

Figure A.1: Topic prevalence score validation



Notes: The x-axis displays the manually annotated values provided by a trained research assistant, while the y-axis reports the GPT-4o mini predictions for the topic prevalence score of the respective objective. Each dot represents an observation for which both human and LLM predictions are available. The validation sample comprises 604 randomly selected five-sentence paragraphs drawn from the full set of speeches. The blue line is the bivariate linear regression line between both variables. The title of each subplot indicates the respective objective. For clarity, values are jittered so that overlapping observations are slightly dispersed.

Figure A.2: Position score validation



Notes: The x-axis displays the manually annotated values provided by a trained research assistant, while the y-axis reports the GPT-5 mini predictions for the topic position score of the respective objective. Each dot represents an observation for which both human and LLM predictions are available. The validation sample comprises 604 randomly selected five-sentence paragraphs drawn from the full set of speeches. Because a position score can only be assigned when a topic is mentioned, the number of observations varies by objective: 116 for price stability, 176 for financial stability, 39 for public debt, 24 for sovereign bond markets, 9 for climate protection, and 10 for distribution. The blue line shows the bivariate linear regression between human and model scores. Subplot titles indicate the respective objectives. For clarity, values are jittered so that overlapping observations are slightly dispersed.

A.2 Prompts

Table A.3: Prompts

Objective	Prompt
Price stability	You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech by a member of the European Central Bank Council. To what extent does the following speech pay attention to the topic of inflation or price stability? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic). In a second step, classify the opinion of the speaker towards the role of central banks to limit the inflation rate and to protect price stability. Return a continuous value with one decimal place from -1.0 to +1.0. The interpretation for the extreme and the central values are: -1.0 is appropriate if the speaker currently sees no risks at all that the inflation rate is or will be too high and that there is therefore currently room for a more expansionary use of conventional or unconventional monetary policy instruments; 0.0 is appropriate if the speaker expresses that the current use of monetary policy instruments is exactly right to reach the central bank's inflation objective; 1.0 is appropriate if the speaker currently sees massive risks that the inflation rate is or will be too high in the future and that there is therefore now the need for a more restrictive monetary policy. The output format should only be: ["Answer 1", "Answer 2"] For Answer 2, return NA if inflation or price stability is not discussed at all.
Financial/banking stability	You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech of a member of the European Central Bank Council. To what extent does the following speech pay attention to the topic of financial market or banking stability? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic). In a second step, classify the opinion of the speaker towards the role of central banks for financial market stability or banking stability. Return a continuous indicator value with one decimal place from -1,0 to +1.0 with the following interpretation for the extreme and the central indicator values: -1.0 is appropriate if the speaker expresses strongly that central banks should use and adjust their instruments to play a more active role to secure financial market stability and banking stability; 0.0 is appropriate if the speaker expresses the opinion that central banks currently play an exactly appropriate role

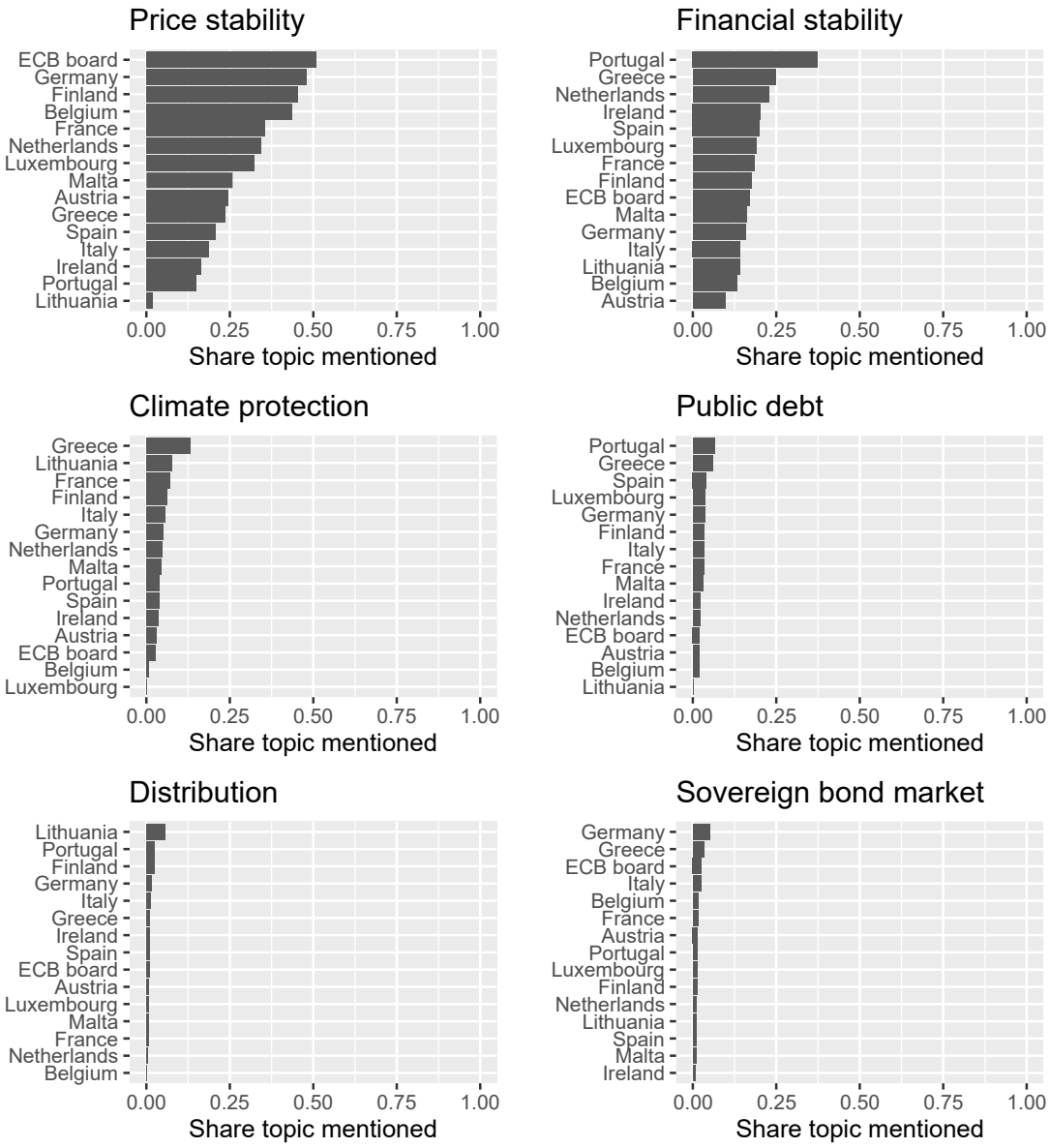
	for financial market stability or banking stability ; 1.0 is appropriate if the speaker is fully convinced that central banks should focus less on financial market stability or banking stability and concentrate instead more on the traditional objectives of monetary policy which is price stability. The output format should only be: ["Answer 1", "Answer 2"]. For Answer 2, return NA if financial market stability or banking stability is not discussed at all.
Public debt	You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech of a member of the European Central Bank Council. To what extent does the following speech pay attention to the topic of public debt or government deficits? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic). In a second step, classify the opinion of the speaker towards the threat of an excessive public debt for the independence of the central bank and the price stability objective of the central bank. Return a continuous value with one decimal place from -1,0 to +1.0 with the following interpretation for the extreme and the central values: -1.0 is appropriate if the speaker expresses strongly that public debt levels are not problematic at all for central bank independence or price stability; 0.0 is appropriate if the speaker has a fully neutral perception of public debt with respect to the risk for central bank independence and price stability; 1.0 is appropriate if the speaker is highly concerned that public debt levels can be problematic for central bank independence or price stability. The output format should only be: ["Answer 1", "Answer 2"]. For Answer 2, return NA if public debt or government deficits are not discussed at all.
Sovereign bond market stability	You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech of a member of the European Central Bank Council. To what extent does the following speech pay attention to the topic of sovereign bond markets (i.e. the market for government bonds)? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic). In a second step, classify the opinion of the speaker towards sovereign spreads (difference in interest rates on national government bonds across Euro states) and if these spreads can be an obstacle for the effectiveness of monetary policy and may obstruct the monetary transmission channel. Return a continuous value with

	one decimal place from -1,0 to +1.0 with the following interpretation for the extreme and the central indicator values: -1.0 is appropriate if the speaker expresses strongly that sovereign spreads are problematic for monetary transmission; 0.0 is appropriate if the speaker has an exactly neutral perception of rising sovereign spreads for monetary transmission; 1.0 is appropriate if the speaker expresses strongly that spreads should not a be a concern for monetary policy and should not distract the central bank from its primary objective of price stability. The output format should only be: ["Answer 1", "Answer 2"]. For Answer 2, return NA if sovereign bond markets are not discussed at all.
Climate protection	You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech of a member of the European Central Bank Council. To what extent does the following speech pay attention to the topic climate change or climate protection? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic). In a second step, classify the opinion of the speaker towards the role of central banks for climate policy. Return a continuous indicator value with one decimal place from -1,0 to +1.0 with the following interpretation for the extreme and the central indicator values: -1.0 is appropriate if the speaker expresses strongly that central banks should use and adjust their instruments to play a more active role in climate policy in the future; 0.0 is appropriate if the speaker expresses the opinion that central banks already support climate policy exactly right; 1.0 is appropriate if the speaker thinks strongly that central banks should focus less on climate policy and concentrate instead more on the traditional objectives of monetary policy like price stability. The output format should only be: ["Answer 1", "Answer 2"]. For Answer 2, return NA if climate change or climate protection is not discussed at all.
Distribution	You have a PhD in Economics and are a researcher specialized in monetary policy. Your task is to classify this speech of a member of the European Central Bank Council. To what extent does the following speech pay attention to the topic of income or wealth distribution? Return a continuous indicator value with one decimal place from 1.0 (no attention to the topic at all) to 5.0 (speech is fully focused on the topic).

	In a second step, classify the opinion of the speaker towards the role of central banks for distribution. Return a continuous indicator value with one decimal place from -1,0 to +1.0 with the following interpretation for the extreme and the central indicator values: -1.0 is appropriate if the speaker expresses strongly that central banks should use and adjust their instruments to play a more active role for distributive policies in the future or should pay more attention to the distributive effects of monetary policy; 0.0 is appropriate if the speaker expresses the opinion that central banks already support distributive policies exactly right; 1.0 is appropriate if the speaker thinks strongly that central banks should focus less on distributive policies and the distributive effects of monetary policy and concentrate instead more on the traditional objectives of monetary policy like price stability. The output format should only be: ["Answer 1", "Answer 2"]. For Answer 2, return NA if income or wealth distribution are not discussed at all.

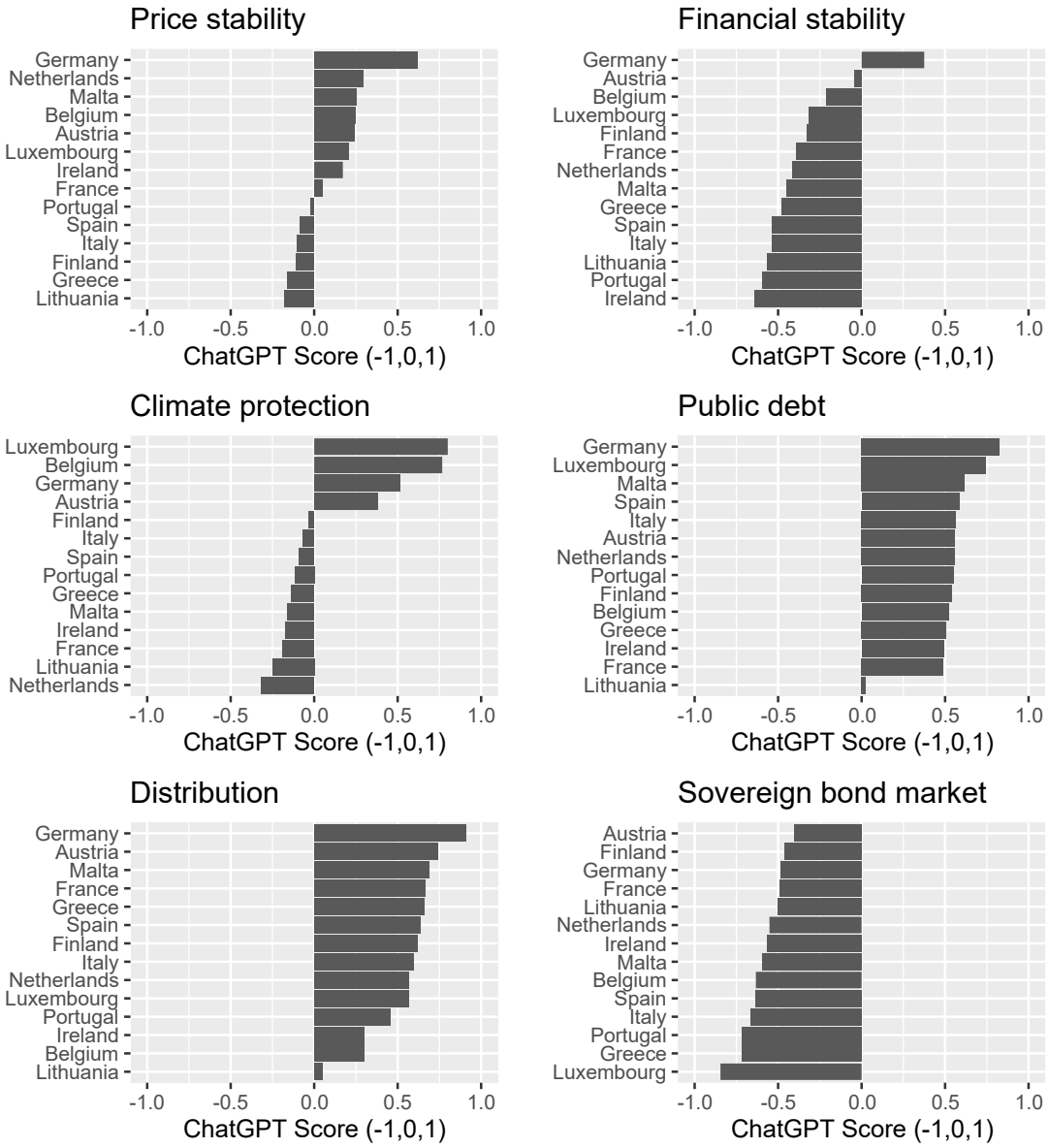
A.3 Additional descriptives

Figure A.3: Dictionary prevalence score by country



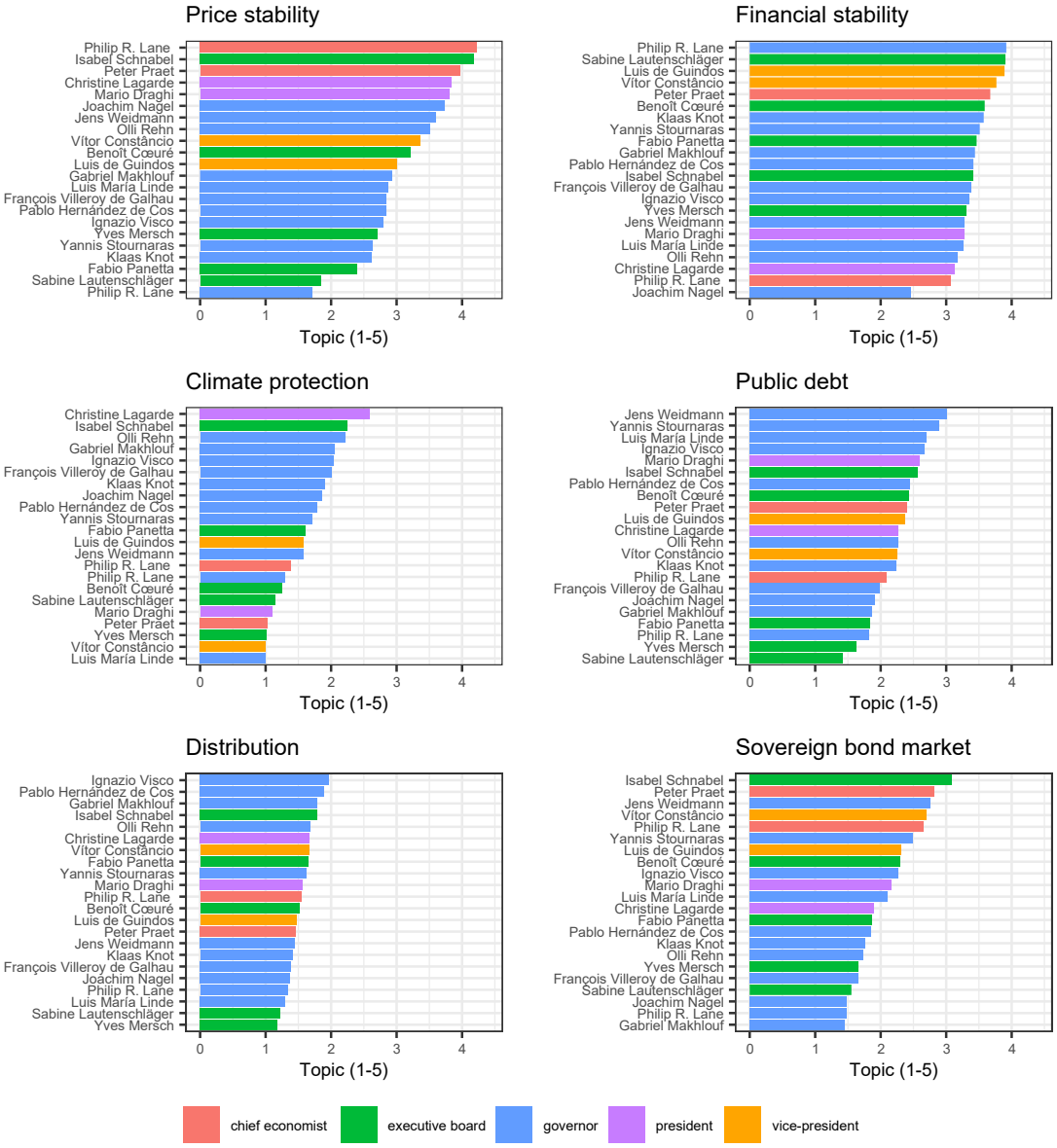
Notes: The x-axis displays the average dictionary terms per country for each objective. Only countries with at least 20 available speeches are included. The title of the subfigures indicates the respective objective.

Figure A.4: LLM topic position score by country



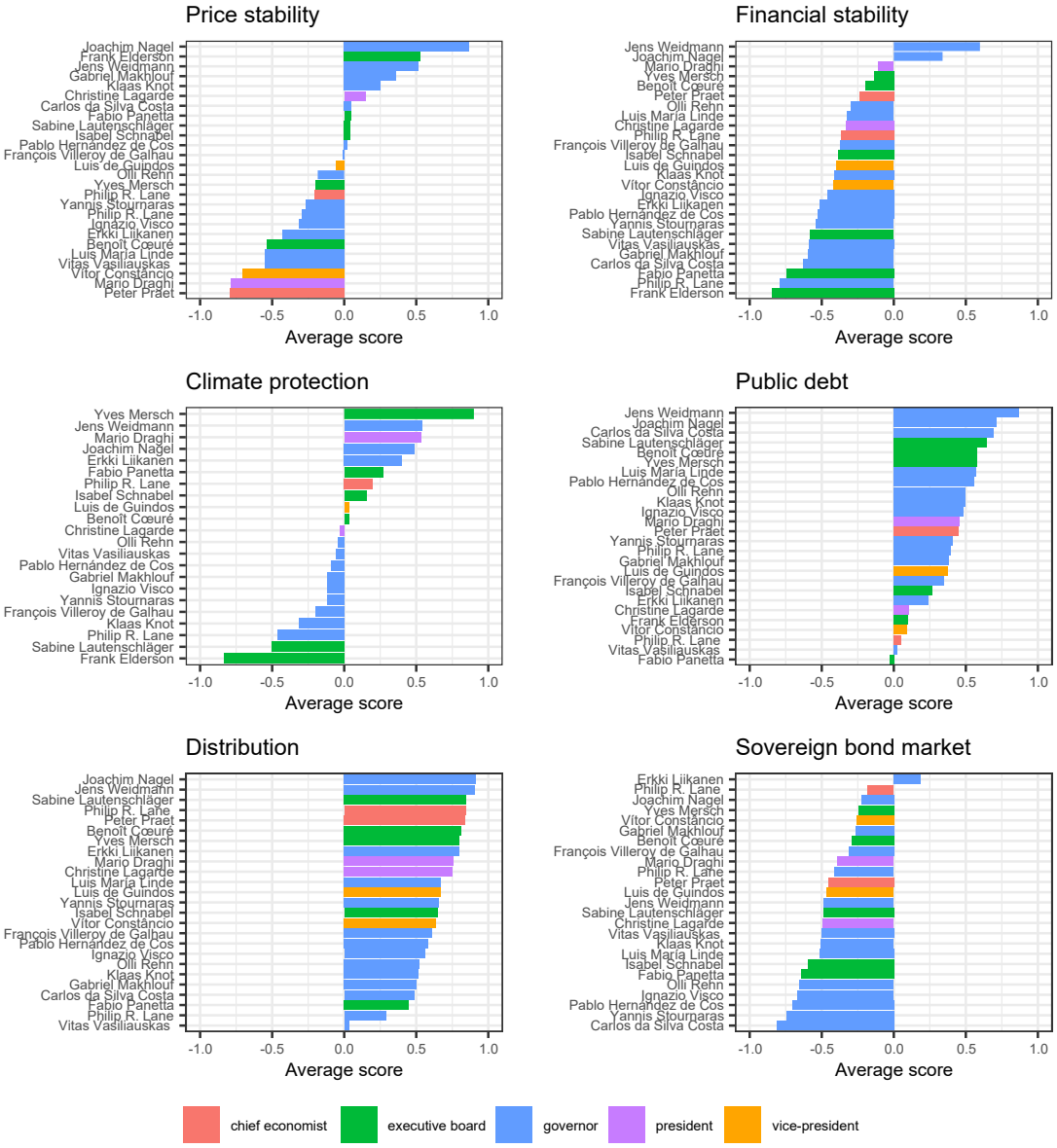
Notes: The x-axis displays the average topic position score per country which ranges from -1 to 1. Only countries with at least 20 available speeches are included. The title of the subfigures indicates the respective objective.

Figure A.5: LLM topic prevalence score by speaker (since 2015)



Notes: On the y-axis are the respective ECB council members. The figure only includes speaker with more than 40 speeches during the years 2015-2024. On the x-axis is the average topic prevalence score per speaker.

Figure A.6: LLM topic position score by speaker (since 2015)



Notes: On the y-axis are the respective ECB council members. Only speaker with more than 20 speeches. On the x-axis is the average topic prevalence score per speaker.

A.4 Additional regressions

Table A.4: Regression Dictionary – Year Fixed Effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00126	0.00012	0.00032	0.00019*	-0.00005	0.00001
	(0.00074)	(0.00041)	(0.00029)	(0.00007)	(0.00005)	(0.00012)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	No
N	2192	2192	2192	2192	2192	2192
R-squared	0.104	0.041	0.090	0.048	0.020	0.062
Cluster SE	Country	Country	Country	Country	Country	Country

Dependent variable: The dependent variable is the topic prevalence score dictionary for the respective objectives.
Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A.5: Regression Dictionary – two-way fixed effects

	Price stability	Financial stability	Climate protection	Public debt	Distribution	Sovereign bond market
Debt/GDP	-0.00158	-0.00020	0.00005	0.00014	-0.00003	-0.00014
	(0.00141)	(0.00101)	(0.00044)	(0.00013)	(0.00007)	(0.00016)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2192	2192	2192	2192	2192	2192
R-squared	0.150	0.074	0.102	0.065	0.027	0.118
Cluster SE	Country	Country	Country	Country	Country	Country

Dependent variable: The dependent variable is the topic prevalence score dictionary for the respective objectives.
Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de