Flipflopping in the European Parliament:

Electoral and Intra-party Politics of MEPs' Migration Speeches*

Tobias Hofmann
Otto Suhr Institute of Political Science
Free University of Berlin
hofmannt@gmail.com

Jim Wagner
Otto Suhr Institute of Political Science
Free University of Berlin
jimvwagner@gmail.com

Abstract

This paper analyzes variation in the position of Members of the European Parliament (MEPs) on asylum-related issues – not only variation between MEPs, their European party groups, and national political parties, but specifically within MEPs across speeches and over time. Prior research has treated MEPs' position as fixed and shown that these positions largely reflect the views expressed by national parties in the domestic political setting and vary along the traditional left-right dimension, but with a clear preference of Euroskeptic MEPs for domestic over European policy solutions. Instead of focusing on variation between politicians and their parties, we are mostly interested in explaining changes in the positions MEPs take over the course of the European election cycle. We investigate the timing of when MEPs portray themselves as being concerned about either the security implications of refugee migration or the human rights of asylum seekers. To empirically test our hypotheses that MEPs adjust their statements on asylum policies in the context of their national, intra-party nomination and elections for the European Parliament (EP), we rely on an artificial intelligence (AI)-based approach to analyzing text as data. Whereas previous research has often used roll-call vote data, manually coded MEP speeches, or automated the analysis of

* Very first draft. We thank OpenAI's Researcher Access Program for their generous support.

heavily preprocessed chunks of words, we employ a large language model to code almost 1,000 entire speeches by over 200 MEPs in EP debates across two election cycles. To highlight the advantages, similarities, and differences of our approach, we compare our results to those produced by human and more traditional computer-based coding, such as *Wordfish*. We discuss alternative measures of inter-coder reliability and the uncertainty in the coding process and explore the potential and implications of prompt engineering for future research into political speech that turns from content to sentiment analysis and the analysis of audience perception, reception, and engagement.

Keywords: European Parliament, election cycle, policy positions, text-as-data, LLM

Flipflopping in the European Parliament:

Electoral and Intra-party Politics of MEPs' Migration Speeches

Actions speaking louder than words

Abraham Lincoln (1856)

Introduction

Actions speak louder than words, but this does not mean that words are irrelevant or that they cannot sometimes provide us with a more detailed and nuanced view of a speaker's preferences, attitudes, or opinions, such as the positions of Members of the European Parliament on a wide range of European Union (EU) policies. While Inger A. Segelström, as a Swedish representative of the Party of European Socialists in the European Parliament and a member of the Swedish Social Democratic Party, may have consistently acted and voted in line with her party group, her plenary speeches show a wide spectrum of views on European asylum policy. In a speech in June 2008, she expressed concerns about the human rights of asylum seekers and the EU

> *moving in the wrong direction and towards a less humane EU. To begin with: to keep people locked up for 18 months is not acceptable. People can go mad, and these people are not felons or criminals but human beings looking for a better life for themselves and their families away from poverty* (Segelström, June 17, 2008).

This stands in stark contrast to her earlier verbal expression of support for common EU rules on asylum, refuge policy, and immigration issues, when she stated in June 2005 that

> *The fact that we can now revise controls and jointly strengthen the external borders means that the EU's borders become more secure. Criminals, drug traffickers, and*

> *human traffickers, who import women and children for the purposes of prostitution, will have fewer opportunities to move freely across borders* (Segelström, June 22, 2005).

Our research question asks why MEPs take the positions they take, but also how we can explain 'flip-flopping' or individual MEPs shifting positions over time and over the course of the European election cycle. Drawing on seminal EU politics work that shows how "MEPs must respond to two different "principals": national parties, who control the selection of candidates in EP elections, and the political groups in the EP, who control a variety of private goods in the EP" (Hix 2002: 688) as well as research on the strategic challenges and effects of two-stage elections in (not just) American politics (Adams & Merrill 2008, Bawn et al. 2012, Brady et al. 2007, Cohen et al. 2009, Saunders & Abramowitz 2004), we, *first*, hypothesize that the MEP agents give in to the sometimes competing pressures from their multiple principles at different points in time. In the run-up and during the years of elections for the EP, MEPs speak more in line with the positions of their national parties. On the other hand, following elections and during non-election years, MEPs have incentives to align their political positions more closely with those of their EP party group. A *second*, partially competing hypothesis focuses on the electoral incentives of national parties, political groups in the EP, and MEPs to 'get their act together' for EP elections (Ferrara & Weishaupt 2004), while they may lack consistent policy positions or a cohesive stance on EU policies between elections.

To empirically test these arguments, we do not rely on legislative roll-call vote data to estimate and track the spatial location of MEPs' political ideal points. While various versions of NOMINATE scores (Poole & Rosenthal 1985, 1991) and similar ideal point estimates have been developed and used to study the EP (Hix et al. 2006, 2007, Lopatin 2013, Salvati 2023) and other legislatures from the U.S. (Clinton et al. 2004) to the United Nations (Voeten 2000), we pursue a text-as-data approach. Ignoring voting actions, we extract MEPs' positions from their words, specifically the positions taken in almost 1,000 EP speeches on asylum policy by over 200 MEPs over the ten-year period 2004-2014. To do so, we employ a large language model (LLM) instead of more conventional computer-based content analyses (e.g., Hopkins & King 2010, Lowe et al.

2010, Proksch & Slapin 2010, Slapin & Proksch 2008) or traditionally hand-coding the text of these speeches.

Following a brief review of the literature on MEP position taking, the development of our own arguments, and a discussion of our empirical strategy, especially the unsupervised AI-coding of our response variable with OpenAI's GPT-4o, we present our regression results on the effects of national parties, European party groups, and various controls on the positions of MEPs on asylum policy before addressing issues of inter-coder reliability and uncertainty in our coding process. We also explore the differences between our coding and previous efforts to code the same corpus of EP speeches (Frid-Nielsen 2018). We conclude with a brief summary and an exploration of the potential and implications of prompt engineering for future, LLM-powered text-as-data research that goes beyond content and sentiment analysis and aims to analyze audience reception, perception, and response to political speech.

Taking MEP Positions

When it comes to the political positions of MEPs, the question mostly boils down to why MEPs cast their votes in support or against proposed European legislation, and the most prominent answers focus on EP-specific institutional structures and party groups as well as MEPs' affiliation with their respective national parties (Hix et al. 2007, Hix & Høyland 2014). This is not to say that personal ideologies and national context do not matter (Cencig & Sabani 2017, Frantescu 2015, Meijers & van der Veer 2019). However, MEPs are not randomly assigned to their national political parties, but select themselves, at least in part, into parties based on ideological and policy considerations, and these parties select themselves in turn into the political groups of the EP. Similarly, national context and public opinion may have a direct impact on MEPs positions, but they also 'trickle down' by shaping the position of national parties as well as national delegations within the EP's political groups (Elomäki et al. 2024).

Of course, these four lines of argument or explanations related to European party groups, national parties, domestic factors, and the personal background and ideologies of MEPs are not mutually exclusive, and they are more nuanced and complicated than they may seem. Simon Hix, Abdul G. Noury, and Gérard Roland (2007) alone identify at least four causal mechanisms or pathways that may explain political group cohesion and party discipline in the EP. One of them focuses on ideological homogeneity among the MEPs of any particular group, where members of the center-left Progressive Alliance of Socialists and Democrats (S&D) share a similar vision of European economic, environmental, or migration policies, just like the Patriots for Europe or members of their Identity and Democracy predecessor are united in their right-wing, populist, and Euroskeptic ideology. A second explanation draws on insights from American politics, suggesting that the leadership of the EP's political groups have strategic incentives to control the legislative agenda so that only legislation with low levels of intra-group controversy or preference heterogeneity come up for a vote. While this agenda setting argument might work well in the context of the U.S. House of Representatives (Cox & McCubbins 2005), the EU's 'rules of the game' that grant the European Commission the exclusive right to initiate European legislation undermine its applicability to the EP and its members. More convincing is the idea that political groups and their leaders can force or entice MEPs with committee assignments, speaking time, and the prospect and promises of leadership positions to strategically toe the group line, even when it clashes with their personal convictions or the policy position of their national parties. Of course, MEPs are not the only players that can act strategically in the legislative game. National parties themselves, with overarching policy goals and coalition building and intra-group, national delegation issue linkage in mind, can pressure their MEPs to fall in line. As a result, they vote with their political group, against their conscience, and possibly even the more short-term policy preferences of the national party.

Despite their differences, all of these alternative causal stories predict individual MEPs to take positions that closely resemble those of their respective European political groups, and additional stories can be told, e.g., about MEPs lacking the necessary resources to make their own, well-informed decisions (Ringe 2009) or the vast majority of EP decisions being largely uncontroversial (Bowler & McElroy), that generally lead to that same conclusions. However,

especially the fourth story can also come with a different ending. National parties that have the ability to force the hands of their MEPs and to enforce EP group cohesion can equally use their influence to pull MEPs away from their European principle and towards their own political position (Meserve et al. 2017). When national party- and European group-positions do not align and MEPs have to choose a side, they know what side their bread is buttered on. It is national political parties that control nominations for elections of the EP, they control the resources required for running what are de facto domestic EP electoral campaigns, and they control the path to a future domestic office. As a consequence, we can expect more national party- than European political group-discipline in the EP. The prediction is that MEPs will, more often than not, vote and express support for policies that please the national party leadership – EP political group cohesion be damned.

Before turning to our own arguments about the shifting influence of national parities and EP groups, it is worth noting that the literature has identified additional variables that can pull the strings of MEPs. They have also been found to respond to national and European public opinion (De Bruycker 2017, Meijers et al. 2019), and it has been argued and shown that their responsiveness may vary between parties and depending on issues and issue salience (Raunio & Wagner 2020, Williams & Spoon 2015). Of course, issue salience may also vary across member states as well as national parties due to country-specific economic, social, and political circumstances. The positions MEPs of the same political group take on agricultural, employment, energy, or migration policy may well be driven by the lobbying efforts of farming interests, levels of unemployment, the domestic energy mix, or member states' geographic location along popular refugee routes.

Changing MEP Positions

We leave it to others to explore how public opinion, issue salience, or national context directly impact or mediate the effects of EP groups and national parties, and empirically, if not intentionally, we control for variation across policy areas by exclusively studying MEPs' positions on the issue of migration. However, we follow in the footsteps of some of this research

by arguing for a conditional effect, where the influence of MEP's national parties and the leaders of their respective political group in the EP is moderated by time. As discussed above, MEPs find themselves under the thumb of two principles, and the literature on their parliamentary behavior suggests that they either strategically vote along the national party line or strategically take the position of their political group. It can be the one or the other, but it can also be both, and the influence of these two principles may systematically wax and wane over the course of the European election cycle. We argue that the carrots and sticks available to the leadership of national parties and EP political groups are more effective and more enticing to MEPs at different points in time. Specifically, MEPs rely on their national principles in the run-up to EP elections when they want to be renominated as candidates and need access to national party resources and active national party support for their election campaigns. Once elected, the power over MEPs shifts to their European principles as they control committee assignments, speaking time, etc. within the EP until the run-up to the next EP elections.

Assuming political survival-maximizing MEPs, who are willing to forego their own ideological convictions and to strategically vote or take on the political position of their national parties and EP groups, depending on their respective power, a systematic shift in the control parties and groups exert over MEPs over time, and holding everything else constant, we hypothesize about the relative policy distance of MEPs from their national parties and EP political groups:

Hypothesis 1: MEPs position themselves more closely to their national party's/EP political group's preferred position before/after elections to the EP.

For us to be able to observe MEPs positions meandering back and forth between the preferred positions of national parties and EP political groups requires them to have different preferred positions. Otherwise, MEPs may zigzag in what is essentially a straight line. Of course, how closely these preferred positions overlap can vary between parties and political groups, across policies, and last, but not least, over time. Research into the success and failure of national parties in EP elections has shown that parties with cohesive EU positions systematically perform

better at the European ballot box than their divided counterparts. Euroskeptic and Europhile voters alike reward parties that take and communicate clear and clearly identifiable positions on the EU (Ferrara & Weishaupt 2004).

Building and expanding on this insight, we argue that the same should hold true for national parties, EP political groups, and their members. If so, they have strong electoral incentives to get their act together around EP elections. Parties and groups agreeing on common policy positions take MEPs on a shorter leash in election years, while allowing them more freedom to speak their mind and express their sincerely held policy positions in between EP elections. In fact, the literature on collective principal action problems and bureaucratic drift (not only) in the EU makes it clear that whenever the ideal points of multiple principals diverge, their ability to reign in renegade agents can be severely limited (e.g., Franchino 2004, Nielson & Tierney 2003, Schneider & Tobin 2013). Therefore, it is safe to assume that the claim that policy disagreement between national parties and EP political groups goes hand in hand with a wider gap between their positions and those of individual MEPs should apply whenever the leaders of European political groups and their constituent national parties fail to form a united policy front. However, focusing on MEP position shifting over the European election cycle, we specifically hypothesize about the absolute policy distance of MEPs from both their national parties and EP political groups:

Hypothesis 2: MEPs position themselves more closely to their national party's and EP political group's preferred position closer to elections to the EP.

Our two hypotheses are fairly general, and we believe that we should be able to observe the two types of position shifting whenever the preferences of MEPs' two principles do not perfectly align. We also believe that we should be able to observe MEPs' strategic positioning in their voting behavior, public statements, social media posts, etc. and across a wide range of policy areas. However, and as discussed in the empirical section of this paper, we specifically focus on position taking through plenary speeches and on the issue of asylum. This does not only allow us to test our own hypotheses, but to replicate and improve on previous research into the predictive

power of national parties' right-left ideological preferences, preferences for or against European integration, and MEPs countries of origin with respect to MEPs' positions during EP election years and the years leading up to the election. It also allows us to assess and validate our novel approach to measuring MEP's political positions.

Measuring MEP Positions

To test our hypotheses on flipflopping in the EP, we study MEPs and their positions over two five-year election cycles or the period 2004-2014. We do so not by relying on roll-call vote data for our dependent variable, but the positions MEPs express in plenary speeches, specifically 876 speeches on asylum policy by 236 MEPs from 19 member states, 62 different national parties, and 5 EP political groups that were previously analyzed by Snorre S. Frid-Nielsen (2018) with the help of Wordfish, a prominent, unsupervised scaling algorithm proposed and developed by Jonathan B. Slapin and Sven-Oliver Proksch (2008). However, unlike Frid-Nielsen (2018), our study utilizes GPT-4o-2024-08-06, which is an advanced OpenAI GPT-4o-family large language model that is optimized for structured output, and while GPT-4o-2024-08-06 is far more capable, we primarily use its capabilities to place MEPs and their speeches along the security dimension of EU asylum policy, i.e., to capture and code our main response variable. Before presenting and discussing that variable, our covariates, regression analysis, and findings, we first detail our AI-based approach, addressing theoretical foundations, data preparation, and speech-coding procedures as well as sketching out our plans for the further evaluation and expansion of this approach for turning strings of text into numeric data.

Our approach is best defined as zero-shot data annotation. Unlike few-shot prompting or chain-of-thought processing, this method does not include examples or demonstrations in the prompt and does not have, like in a chat-style conversation, multiple instances of the LLM reason over the task. Zero-shot data annotation with LLMs has proven highly accurate, outperforming s Amazon Mechanical Turk (MTurk) crowd workers by 25% on average (Gilardi 2023). Unlike few-shot prompting, this method does not include examples or demonstrations in the prompt. We leverage the LLM's superior ability to understand context and nuance in political

speech, allowing us to grasp the subtleties of security-related rhetoric in a holistic manner. As long as we provide the model with role, instruction, task, and context prompts, we assume that the LLM's inherent contextual understanding, generalization, reasoning, and knowledge capabilities are sufficient to operate solely based on our prompts. These capabilities arise from a nuanced interplay between memorization and generalization, coupled with adequate task-relevant pretraining data (Antoniades et al. 2024). The primary advantage of an LLM in this context is its generalization capability, eliminating the need for highly specialized models. Furthermore, the persistent lack of evaluations in the field and strong prompting guidelines lead us to a more bare-bone approach that relies on the LLM's general capabilities to answer intuitive questions based on the same linguistic setup that human coders are expected to work with and employ.

A key strength of LLMs lies in their capacity to bridge qualitative and quantitative research methodologies. By translating nuanced, natural language questions into precise quantitative measures, LLMs enable social scientists to seamlessly integrate their domain expertise into computational frameworks. While this capability can facilitate multidimensional analyses and open new avenues for interdisciplinary collaboration (Ziems et al. 2024), our paper focuses on a single dimension of interest to political scientists – the position of legislators on the security dimension of asylum policy. Our main objective is to discretize the text of MEP speeches into the dependent variable of our hypotheses. Once that variable is measured and by documenting and evaluating the various stages of the analysis, we can also gain some initial insights into the similarities and differences, methodological advantages and disadvantages of our LLM-based zero-shot text discretization approach vis-à-vis more traditional quantitative text analysis and the hand-coding of MEPs' speeches and their asylum-related security stance. Such a comparison is necessary to assess whether we can make a methodological contribution or to evaluate our contribution to the flourishing text-as-data literature that, in the EU studies-context, has explored everything from special interest group lobbying (Klüver 2009, 2011) and deliberations in the Council of the European Union (Wratil & Hobolt 2019) to party positions (Hjorth et al. 2015) and position taking in EP speeches (Proksch & Slapin 2010).

As mentioned before, we use our AI-approach to code 876 speeches "on asylum in the period 2004-2014 were extracted, based on debate topics and a keyword search for asylum" (Frid-Nielsen 2018: 349). These speeches were selected for reasons of convenience as well as to enable us to conduct a direct comparison with the scores produced by Wordfish and human coding. It is this second reason why we do not report results of an 'upstream' step of our analysis, where we made use of GPT-4o-2024-08-06's capabilities to classify MEP speeches and identify asylum-related speeches among all plenary speeches delivered over the course of more than two election cycles. However, when asking the model to assess the 876 speeches on the degree they address six issue areas (economy, healthcare, education, environment, security, and migration), the vast majority received a high migration rating.

Unlike the data preparation for most quantitative text analyses, our analysis does not require extensive preprocessing of the corpus. It does not include the removal of stop words, stemming, etc. as the LLM processes natural language more like a qualitative, human coder, following human-readable coding instructions. However, we checked for spelling mistakes in the transcribed speeches and such occasional annotations as "(applause)" that are not part of the speech, but indicate audience response. Otherwise, we only needed to assign each speech a unique identifier as well as speaker IDs. Frid-Nielsen (2018) lumped speeches together at the national party level. This is in line with the common recommendations to aggregate input documents into larger 'bags of words' (Hjorth et al. 2015, Proksch & Slapin 2010), but is not required for our approach, allowing us to analyze intra-MEP variation over time. At the same time, it does not prevent us from collapsing individual speech scores over MEPs, national parties, or political groups after coding. While our main analysis was conducted with all 876 speeches, we did initial tests and parameter evaluations, using a fixed seed (seed=42) for reproducibility, with a randomly sampled subset of 100 speeches.

For the final coding of all MEP speeches, we utilized GPT-4o-2024-08-06 via the OpenAI API and set the temperature and max. tokens parameters to 0.4 and 1,000, respectively. Temperature controls the randomness of the model's responses, with lower values producing more deterministic results. Perhaps somewhat counterintuitively, we found that a temperature of

0.4 produced the lowest variance in results across multiple runs, making it the best choice for ensuring consistent and reliable results for our specific task of rating asylum speeches on the security dimension. The maximum number of tokens was set to 1,000 as responses are not expected to exceed this number. With these parameters set, we designed prompts that simulate the instructions provided to human coders in a manual content analysis setting (Halterman & Keith 2024). This meant that we sent each speech to the model with variations of the following components:

- *Role:* trained political science coder
- *Instruction:* rate the stance of the speaker on the security issue dimension using a scale from 1 to 10
- *Task:* decode the speech
- *Context:* full text of the speech

For the main analysis, these components were implemented in the following prompts:

- *System prompt:* "You are a trained political science coder tasked with rating the stance of the speaker on the security issue dimension. Return only a number between 1 and 10. A 10 indicates the speaker is completely in favor of the security issue, and a 1 indicates the speaker is completely against the security issue."
- *User prompt:* "Decode this speech: {speech}"

We employed the instructor library for structured output parsing, defining a data model:

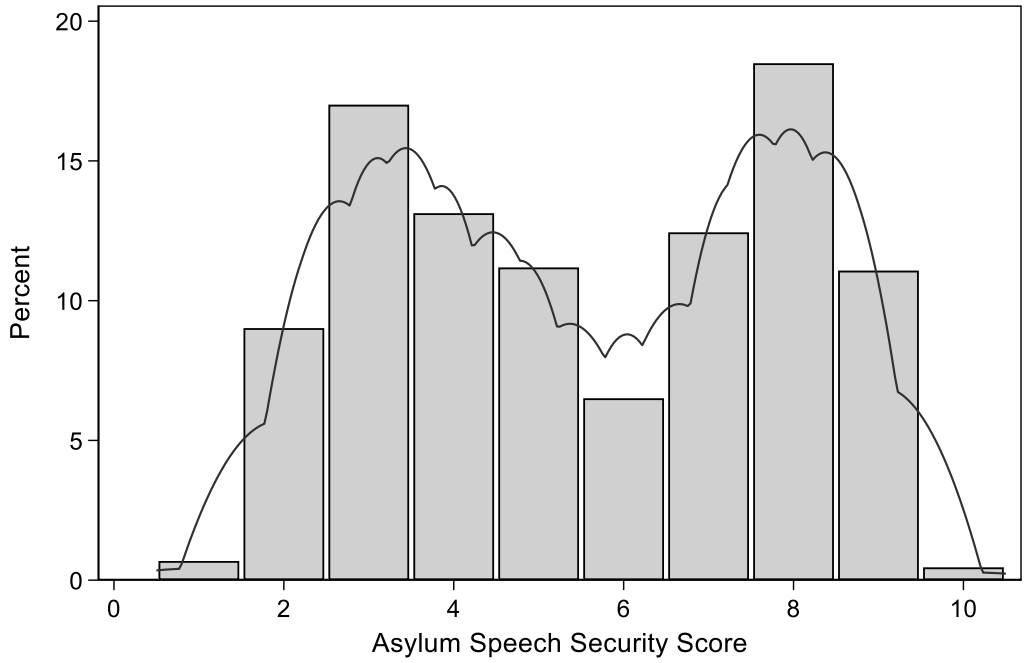
- class SecurityStance(BaseModel): score: float

Practically, this setup allowed us to automatically extract the numeric ratings and append them to our dataset as a new 'security dimension' column. Automating the structured output parsing substantially increases the iteration speed, streamlining data processing and enabling us to conduct ad-hoc prompt design without the need for extensive instruction-tuning on the expected output format. By defining the specific SecurityStance data model with a structured score field, we avoid such potential issues as inconsistent formatting or misinterpretation of the

LLM's response. This approach enhances the reliability and efficiency of the coding process, simplifying the analysis itself as well as the comparison across multiple speeches.

Figure 1 visualizes the distribution of this new *SecurityStance* variable. The histogram shows a clear bimodal distribution, indicating that MEP's asylum speeches and the underlying security issue are ideologically charged.

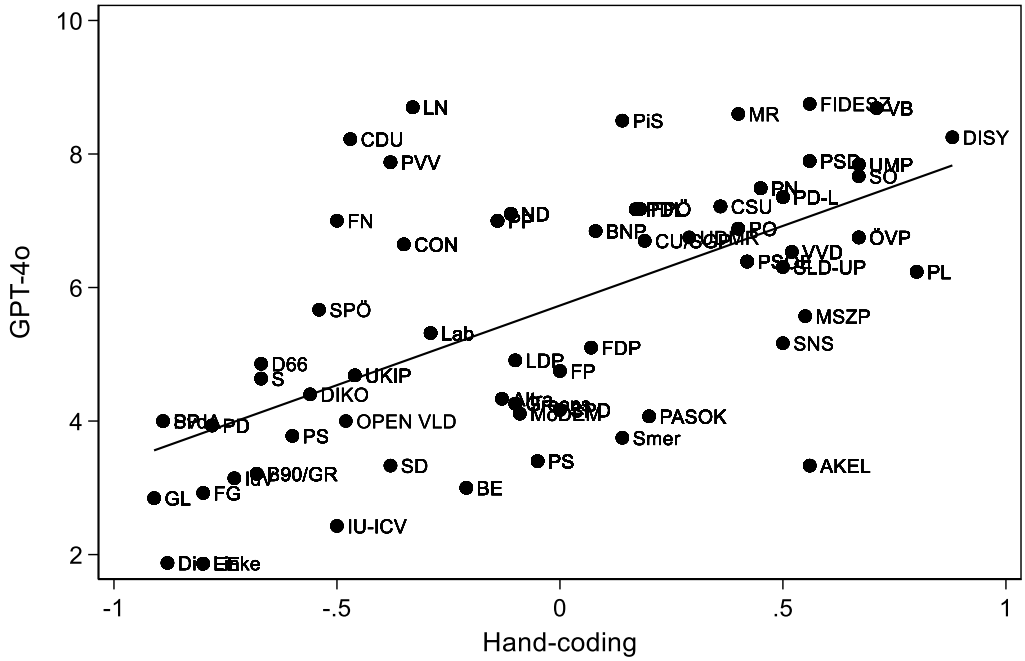
Figure 1: SecurityStance of MEP's Asylum Speeches



Our evaluation of the instructor library demonstrated its superior performance in both extracting 'correct' numeric ratings and reducing output variance compared to manual parsing methods. In short, it combines the advantages of speed, consistency, etc. of quantitative text analysis, while using transparent coding instructions that are almost identical to those used in the context of large expert-based, hand-coding endeavors, such as the Comparative Manifestos Project (Budge et al. 2001, Klingemann et al. 2006, Volkens et al. 2013).

With respect to the evaluation and validation of our model and data, we undertook a multi-step, only partially sequential process. This process included, first, parameter testing, i.e., assessing the effects of different temperature settings on rating consistency. Second, we assessed the quasi-intercoder reliability by rating each speech multiple times with identical prompts. With the system prompt mentioned above, we archived a Cronbach's Alpha of .844 across multiple runs, using disaggregated data, and of .989, when aggregating individual scores at the national party level to make them comparable to those generated with the help of Wordfish and by hand-coding a single, randomly selected speech per national party (Frid-Nielsen 2018). Third, we correlated our SecurityStance variable with the Wordfish-generated and hand-coded scores mentioned in the previous step, yielding consistently strong and highly statistically significant relations between the manual, traditionally automat, and AI coding methods (Pearson correlation coefficient: .548 to .771, $p < .001$). Figure 2 depicts the relationship between the hand-coded and our SecurityStance variable. National parties above the regression line, such as the Christian Democratic Union of Germany (CDU), have a lower hand-coded security score than our analysis of all asylum speeches of CDU MEP suggests. On the other hand, our model underestimates or human coding overestimates the security focus of Cyprus' Progressive Party of Working People (AKEL). We still need to analyze whether there are any systematic patterns to where human, Wordfish and LLM codes deviate. Identifying such error patterns should be able to inform future improvements to our approach, especially improvements related to the system prompt.

Figure 2: Hand-coded vs. AI-generated SecurityStance



In a fourth step, we repeated the second and third step for differently worded system prompts in line with generic prompt engineering recommendations (White 2023). To assess the quality of the data generated by different prompts, we relied on three fundamentally different approaches. Statistically, we assessed the variance in the coding of individual speeches over multiple runs with the same prompt and across different system prompts. The other two approaches can only be implemented in an AI-context. Here, we asked the model to provide a structured, numeric evaluation of its own coding performance for each speech. In essence, GPT-4o-2024-08-06 provided us with its own estimate of uncertainty in the coding process or its confidence that the SecurityStance for a specific MEP speech is correct. The final approach consisted of requesting and reading unstructured coding output, i.e., having the model provide an explanation for why a given speech was decoded in a particular way. This is similar to asking human experts to explain their coding decisions and justify why they chose a specific score.

Of course, all of these are only initial steps towards ensuring the reliability and validity of our coding approach and will be further refined. In addition, we have plans to conduct several additional types of evaluations and to develop extensions to our model. On the refinements side, we want to focus more on prompt engineering and to more systematically experiment with different instructions that will optimize LLM performance. Possible extensions include moving beyond a single SecurityStance code per speech and treating the security aspect of the EU's asylum policy as a multifaceted and multidimensional issue. This would bring our coding process more in line with human coding approach used by the beforementioned Comparative Manifestos Project, which distinguishes between such categories as burden sharing, readmission procedure, Schengen, external border, human rights, and freedom or fifteen codes that "reflect both the substantial and functional aspects of the asylum policy dimension" (Frid-Nielsen 2018). Another step would be moving beyond asylum policy altogether. After all, analyzing the 876 asylum speeches was partially driven by convenience. There is no reason not to expand the analysis across policies and time, i.e., to include other contentious issues, such as the environment or monetary policy, and speeches from before 2004 and after 2014.

Analyzing MEP Positions

Before we can test our hypotheses as well as try to replicate the results of Frid-Nielsen (2018), replacing the Wordfish-coded response variable with our own SecurityStance variable, we need to quickly discuss the core covariates, other independent variables, and controls of the analysis. The independent variables needed for the replication include a binary variable that indicates if a MEP and his national party are from a new Central or Eastern European member state, which joined the EU in 2004 or later. Two additional variables from Frid-Nielsen (2018) are a party manifesto-based measure of national parties' left-right ideological position as well as their stance towards European integration. Controls that only form part of some of our regression models are various sets of dummy variables that allow us to estimate member state, EP political group, and national party effects.

What about the preferred position on asylum policy of MEP's national parties and EP political groups? Admittedly, this is one of the major weaknesses of our study. We simply calculate the average SecurityStance score across all the speeches of MEPs from a given political group and party. In the future, we will replace these measures with data from Chapel Hill Expert Surveys on national party positioning on ideology and policy issues across all EU member states (Jolly et al. 2022), data on party policies that are designed to return asylum seekers and immigrants to their countries of origin (Benoit & Laver 2006), as well as manifesto-derived political position data. For now, this gives us 62 unique national party- and 5 EP political group-values that are constant over time. Table 1 shows the SecurityStance score and their standard deviation for all groups in the EP and the values for a few national parties. Despite their shortcomings, they at least pass a plausibility test, with the conservative and more right-wing European People's Party (PPE) and Europe of Freedom and Democracy (EFD) groups scoring higher than left, liberal, and green EP groups. Similarly, it is not surprising and in line with expectations that the center-right CDU has a stronger security focus than the far-left Die Linke and Germany's green party, Alliance 90/The Greens (B90/GR).

Table 1: Group and Party SecurityStance

EP Political Groups			**Selected National Parties**	
	Mean	*Standard Deviation*		*Mean*
ALDE	4.81	1.99	AKEL	3.42
EFD	6.99	2.00	Altra	3.83
G/EFA	3.40	1.31	B90/GR	3.58
GUE/NGL	3.57	1.65	BE	3.41
NI	7.28	2.12	BNP	6.23
PPE	6.92	1.81	CDU	8.00
SD	4.60	2.03	CON	6.47
			CSU	6.50
			CU/SGP	6.10
			D66	3.86
			DIKO	4.20
			DISY	8.25
			Die Linke	2.50
			…	…

As the response variables are continuous, we can use ordinary least square to estimate the various regression models. In a first step, we reproduce the results presented in table 2 of Frid-Nielsen (2018: 356), specifically the multiple regression model M4. Our replication results in Model 1 of Table 2 are identical, showing a positive and statistically significant effect of a national party's stance on EU integration and its geographical location in Central and Eastern Europe on the position of its MEPs. When replacing the Wordfish-generated dependent variable with ours, things change. Independent of whether we run the regression for the 62 national parties (Model 2), 876 speeches (Model 3), or 876 speeches including time effects (Model 4), we find different, but highly consistent results for the right-left and EU integration variables. Only the geographic variable or age of membership has the same positive and statistically significant impact. At the moment, we do not have a good explanation for these similarities and differences.

Table 2: Replication

	Model 1	**Model 2**	**Model 3**	**Model 4**
	Frid-Nielsen (2018)	*SecurityStance*	*SecurityStance*	*SecurityStance*
Right-left	-0.0029	0.043**	0.062***	0.063***
	(0.008)	(0.016)	(0.006)	(0.006)
EU integration	0.015**	-0.015	-0.014***	-0.018***
	(0.006)	(0.011)	(0.004)	(0.004)
Old/new member	1.059***	0.963*	0.853***	0.839***
	(0.294)	(0.538)	(0.228)	(0.228)
Constant	-0.243*	5.572***	5.834***	5.734***
	(0.134)	(0.246)	(0.088)	(0.700)
Year effects	no	no	no	yes
N	62	62	876	876
R^2	0.272	0.150	0.113	0.136

Standard errors in parentheses; *p < 0.1, **p > 0.05, and ***p > .01 (two-tailed).

When it comes to our own hypotheses, there are several ways to test whether national parties and EP political groups influence MEPs expressed policy positions, if those positions change over time, and if they shift in the expected direction. To recall, our first hypothesis predicts that MEPs position themselves more closely to their national party's/EP political group's preferred position before/after elections to the EP. For now, we pursue a similar approach as in Models 3 and 4 of Table 2, i.e., we run OLS regressions with individual speeches as the unit of analysis and the response variable SecurityStance. What changes are the covariates. Instead of variables that measure general national party attributes, we now include national party/EP political group SecurityStance, an inverse election counter that starts in the year following an election for the EP, and the interaction effect between this time variable and national parties and EP political groups' respective SecurityStance. Alternatively, we also specify the model with a binary election year variable instead of the continuous election cycle variable. Finally, we include MEP fixed effects in all models.

Table 3: Party and Group Effects

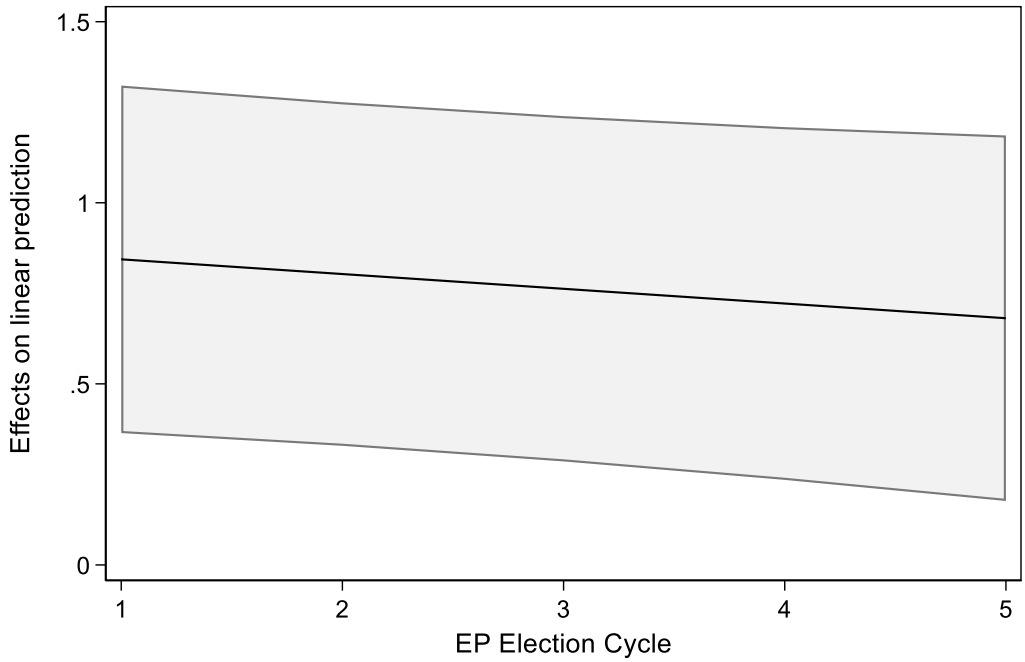
	Model 5	Model 6	Model 7	Model 8
National party position	0.885***		0.801***	
	(0.251)		(0.242)	
Political group position		0.936***		0.862***
		(0.271)		(0.261)
Interaction effect	-0.041	-0.038	-0.085	-0.080
	(0.031)	(0.036)	(0.111)	(0.125)
Election cycle	0.188	0.165		
	(0.177)	(0.199)		
Election year			0.387	0.356
			(0.617)	(0.682)
Constant	0.988	0.974	1.361	1.295
	(1.630)	(1.661)	(1.588)	(1.608)
MEP effects	yes	yes	yes	yes
Observations	876	876	876	876
R-squared	0.658	0.658	0.657	0.657

Standard errors in parentheses; *p < 0.1, **p > 0.05, and ***p > .01 (two-tailed).

When it comes to the question whether national parties and EP political groups influence MEPs, the answer is a resounding yes. All national party and political group position coefficients are positive and highly statistically significant. Of course, that is to be expected. It is in line with the discussions in this paper's literature and theory section. Furthermore, the way these covariates were created makes it virtually impossible for the coefficients to be anything but positive and significant. However, our first hypothesis also claims that the influence should change in opposing directions over time. That is not what we find. Not only are our election cycle and election year variables not statistically significant, but if we inspect the interaction effects, we find that – were they significant – the influence of both principles declines as EP elections come closer. The marginal effects plot in Figure 3 illustrates this for MEPs national parties. The confidence intervals around the average marginal effects on the left and the right, in years 1 and 5 of the election cycle overlap. If they did not overlap, the plot would tell us that the control national parties have over their MEPs steadily declines following an EP election. This is the opposite of what we would expect. Instead, we hypothesized such a decline for the power of

EP political groups, where the marginal effects point in the right direction, but are also not statistically significant.

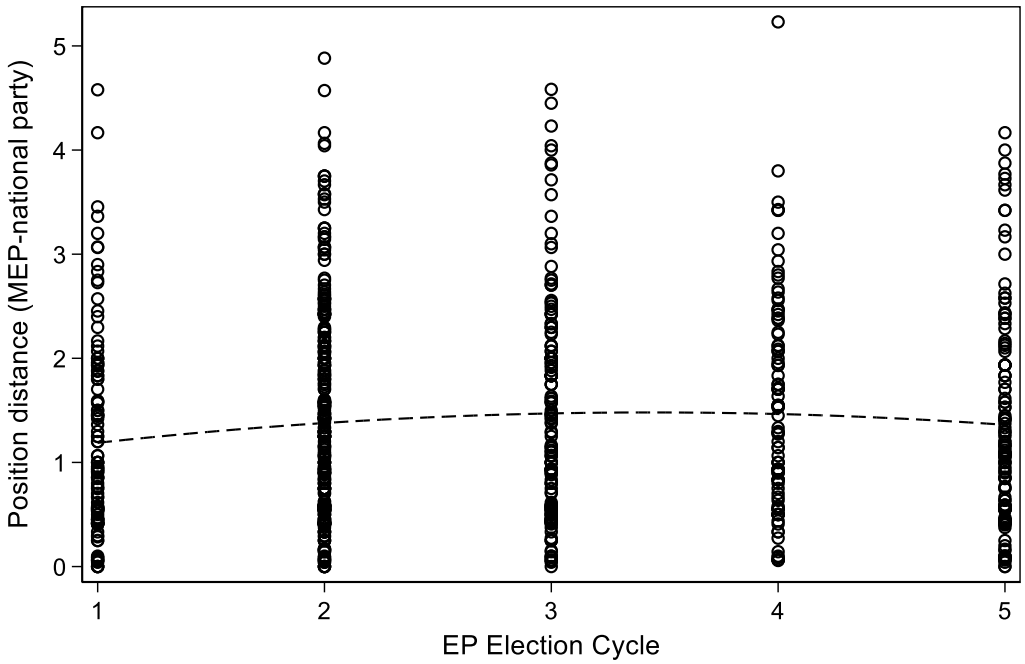
Figure 3: Marginal Effect of National Party Positions over Time



The second hypothesis expects MEPs to position themselves more closely to their respective national party's as well as their EP political group's preferred position closer to elections to the EP, i.e., shortly before and after the elections, but not in the middle of the election cycle. When we test this claim, our regression results are once again encouraging and disappointing at the same time. While the coefficients point in the right direction, the results are unfortunately not statistically significant. When closely inspecting Figure 4, one can see a slight bend in the line of best fit, showing a (non-significant) inversely u-shaped relationship between how far MEPs' SecurityStance deviate from that of their national party principals and the EP election counter. Replacing national party positions with those of MEPs' European principles yields virtually identical results. A positive election counter coefficient indicates that MEPs initially move away from the preferred European asylum policies of their national parties and

MEP political groups before the gravitational pull of a negative squared coefficient brings them back towards the end of the election cycle.

Figure 4: Absolute Distance from National Party Positions over Time



To summarize our very preliminary findings, the currently available data rejects both of our hypotheses. MEPs respond positively to national party and EP political group pressure, keeping their own political positions on European asylum policy in sync with those of their political overlords. Leaving firm statistical ground behind and starting to speculate, our findings might hint towards a slight advantage for the leadership of MEP's political groups in the EP. The position coefficients are marginally larger in Models 6 and 8 than 5 and 7. Furthermore, the interaction effects are larger in Models 5 and 7 than 6 and 8. Because they are not statistically significant, this does not really mean anything. However, it might hint at MEPs not only speaking more freely on the topic of asylum in the run-up to EP elections, but distancing and distinguishing themselves slightly more from the preferred position of their national party than that of the EP political group. At the same time, the non-significant results for our second

hypothesis might also hint at MEPs enjoying the greatest amount of freedom in between EP elections and feel the pressure and the need to fall in line with the positions of their national and European principals when it is electorally important to get their act together and to lead voters into believing that they all share a common policy position. Again, this is mere speculation and not currently supported by our data. Of course, if it was true, it would warrant further theoretical and empirical inquiry into the possible election-strategic or other reasons for such a differentiated shift in speaking behavior.

Conclusion

The purpose of this paper is to make a substantive and methodological contribution to EU studies and the text-as-data literature. Following a brief discussion of prominent arguments and studies on the voting and position taking of MEPs, we developed two competing theoretical arguments that MEPs face different pressures before, during, and after EP elections. Our first hypothesis claims that in the run-up to elections, national factors are more important. MEPs care about domestic public opinion, voter preferences, and especially getting renominated and receiving campaign support from their national party. Accordingly, we expect MEPs to strategically adjust what they say in plenary speeches in the directions of national party leaders preferred political positions. Once (re)elected, the calculus shifts, and MEPs try to appease and appeal to EP political group leaders. The second hypothesis argues that what overrides these considerations is the need of national parties, EP political groups, and MEPs to take a united policy stance around EP elections. While we still find both of these arguments fascinating, our empirical analysis of 876 speeches does not lend any meaningful support to the idea that MEPs strategically flipflop their way through the EP's electoral cycle. Both of our hypotheses are rejected by the data on MEPs' asylum policy positions in the years 2004-2014.

This leaves us with the paper's methodological contribution. Here, the goal is to contribute to the AI revolution in the social sciences by using the analytical capabilities to meaningfully extract numbers from text in an innovative way that straddles and combines the best aspects of qualitative hand-coding and quantitative text analysis. We provide a brief discussion of LLMs

before digging into our process of coding MEP speeches with the help of OpenAI's GPT-4o-2024-08-06. We also go into detail of how we and one can assess the quality of AI-extracted data and compare and evaluate our approach vis-à-vis human coding and traditional quantitative text-as-data approaches. While conscious of the limitations and model biases inherent in LLMs and very much aware of the preliminary nature of our ongoing work and results, we strongly believe that our approach offers new possibilities for analyzing political text and has the opportunity to develop into a robust, reliable method for automated political text analysis. Of course, using LLMs to code the political positions of MEPs from EP speeches is not the end of our AI journey. In the future, we are interested in expanding the approach to study issues related to audience perception, reception, and engagement, possibly with the help of and including synthetic samples for quasi-survey experiments.

Bibliography

Adams, James, and Samuel Merrill III. "Candidate and party strategies in two-stage elections beginning with a primary." American Journal of Political Science 52.2 (2008): 344-359.

Antoniades, Antonis, Xinyi Wang, Yanai Elazar, Alfonso Amayuelas, Alon Albalak, Kexun Zhang, and William Yang Wang. "Generalization vs Memorization: Tracing Language Models' Capabilities Back to Pretraining Data." arXiv 2407.14985 (2024): 1-12.

Bawn, Kathleen, Martin Cohen, David Karol, Seth Masket, Hans Noel, and John Zaller. "A theory of political parties: Groups, policy demands and nominations in American politics." Perspectives on Politics 10.3 (2012): 571-597.

Benoit, Kenneth, and Michael Laver. Party policy in modern democracies. Routledge, 2006.

Bowler, Shaun, and Gail McElroy. "Political group cohesion and 'hurrah' voting in the European Parliament." Journal of European Public Policy 22.9 (2015): 1355-1365.

Brady, David W., Hahrie Han, and Jeremy C. Pope. "Primary elections and candidate ideology: Out of step with the primary electorate?" Legislative Studies Quarterly 32.1 (2007): 79-105.

Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, and Eric Tanenbaum. Mapping policy preferences: estimates for parties, electors, and governments, 1945-1998. Oxford University Press, 2001.

Cencig, Elisa, and Laura Sabani. "Voting behaviour in the European Parliament and economic governance reform: does nationality matter?" Open Economies Review 28 (2017): 967-987.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. "The statistical analysis of roll call data." American Political Science Review 98.2 (2004): 355-370.

Cohen, Marty, et al. The party decides: Presidential nominations before and after reform. University of Chicago Press, 2009.

Cox, Gary W., and Mathew D. McCubbins. Setting the Agenda: Responsible Party Government in the US House of Representatives. Cambridge University Press, 2005.

De Bruycker, Iskander. "Politicization and the public interest: When do the elites in Brussels address public interests in EU policy debates?" European Union Politics 18.4 (2017): 603-619.

Elomäki, Anna, Johanna Kantola, Petra Ahrens, Valentine Berthet, Barbara Gaweda, and Cherry Miller. "The role of national delegations in the politics of the European Parliament." West European Politics 47.6 (2024): 1251-1275.

Ferrara, Federico, and J. Timo Weishaupt. "Get your act together: Party performance in European Parliament elections." European Union Politics 5.3 (2004): 283-306.

Franchino, Fabio. "Delegating Powers in the European Community." British Journal of Political Science 34 (2004): 269-293.

Frantescu, Doru Petrisor. "Values topple nationality in the European Parliament." European View 14.1 (2015): 101-110.

Frid-Nielsen, Snorre Sylvester. "Human rights or security? Positions on asylum in European Parliament speeches." European union politics 19.2 (2018): 344-362.

Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli. "ChatGPT outperforms crowd workers for text-annotation tasks." Proceedings of the National Academy of Sciences 120.30 (2023): 1-3.

Halterman, Andrew, and Katherine A. Keith. "Codebook LLMs: Adapting Political Science Codebooks for LLM Use and Adapting LLMs to Follow Codebooks." arXiv 2407.10747 (2024): 1-29.

Hjorth, Frederik, Robert Klemmensen, Sara Hobolt, Martin Ejnar Hansen, and Peter Kurrild-Klitgaard. "Computers, coders, and voters: Comparing automated methods for estimating party positions." Research & Politics 2.2 (2015): 1-9.

Hix, Simon. "Parliamentary behavior with two principals: Preferences, parties, and voting in the European Parliament." American Journal of Political Science (2002): 688-698.

Hix, Simon, and Bjørn Høyland. “Political Behaviour in the European Parliament.” Shane Martin, Thomas Saalfeld, and Kaare W. Strøm (eds). The Oxford Handbook of Legislative Studies. Oxford University Press, 2014.

Hix, Simon, Abdul Noury, and Gerard Roland. "Dimensions of politics in the European Parliament." American Journal of Political Science 50.2 (2006): 494-520.

Hix, Simon, Abdul G. Noury, and Gérard Roland. Democratic politics in the European Parliament. Cambridge University Press, 2007.

Hopkins, Daniel J., and Gary King. "A method of automated nonparametric content analysis for social science." American Journal of Political Science 54.1 (2010): 229-247.

Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. "Chapel Hill Expert Survey Trend File, 1999-2019. " Electoral Studies 75 (2022).

Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Ian Budge, and Michael D. McDonald. Mapping policy preferences II: estimates for parties, electors, and governments in Eastern Europe, European Union, and OECD 1990-2003. Oxford University Press, 2006.

Klüver, Heike. "Measuring interest group influence using quantitative text analysis." European Union Politics 10.4 (2009): 535-549.

Klüver, Heike. "The contextual nature of lobbying: Explaining lobbying success in the European Union." European Union Politics 12.4 (2011): 483-506.

Lopatin, Esther. "The Changing Position of the European Parliament on Irregular Migration and Asylum under Co-decision." *JCMS: Journal of Common Market Studies* 51.4 (2013): 740-755.

Lowe, Will, Kenneth Benoit, Slava Mikhaylov, and Michael Laver. "Scaling policy preferences from coded political texts." Legislative studies quarterly 36.1 (2011): 123-155.

Meijers, Maurits J., Christina J. Schneider, and Asya Zhelyazkova. "Dimensions of input responsiveness in the EU: actors, publics, venues." Journal of European Public Policy 26.11 (2019): 1724-1736.

Meijers, Maurits J., and Harmen van der Veer. "MEP responses to democratic backsliding in Hungary and Poland. An analysis of agenda-setting and voting behaviour." JCMS: Journal of Common Market Studies 57.4 (2019): 838-856.

Meserve, Stephen, Joseph Robbins, and Frank Thames. "Multiple principals and legislative cohesion." Legislative Studies Quarterly 42.4 (2017): 515-548.

Nielson, Daniel L., and Michael J. Tierney. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." International Organization 57 (2003): 241-276.

Poole, Keith T., and Howard Rosenthal. "A spatial model for legislative roll call analysis." American journal of political science (1985): 357-384.

Poole, Keith T., and Howard Rosenthal. "Patterns of congressional voting." American journal of political science (1991): 228-278.

Proksch, Sven-Oliver, and Jonathan B. Slapin. "Position taking in European Parliament speeches." British Journal of Political Science 40.3 (2010): 587-611.

Raunio, Tapio, and Wolfgang Wagner. "Party politics or (supra-) national interest? External relations votes in the European Parliament." Foreign Policy Analysis 16.4 (2020): 547-564.

Ringe, Nils. Who decides, and how? Preferences, uncertainty, and policy choice in the European Parliament. Oxford University Press, 2009.

Salvati, Eugenio. "Debating Legitimacy and Solidarity in the European Parliament: Patterns of Opposition and Conflict During the Pandemic." JCMS: Journal of Common Market Studies (2023).

Schneider, Christina J., and Jennifer L. Tobin. "Interest Coalitions and Multilateral Aid Allocation in the European Union." International Studies Quarterly 57 (2013): 103-114.

Sebők, Miklós, Sven-Oliver Proksch, Christian Rauh, Péter Visnovitz, Gergő Balázs, and Jan Schwalbach. "Comparative European legislative research in the age of large-scale computational text analysis: A review article." International Political Science Review (2023): 1-22.

Slapin, Jonathan B., and Sven-Oliver Proksch. "A scaling model for estimating time-series party positions from texts." American Journal of Political Science 52.3 (2008): 705-722.

Saunders, Kyle L., and Alan I. Abramowitz. "Ideological realignment and active partisans in the American electorate." American Politics Research 32.3 (2004): 285-309.

Voeten, Erik. "Clashes in the Assembly." International organization 54.2 (2000): 185-215.

Volkens, Andrea, Judith Bara, Ian Budge, Michael D. McDonald, and Hans-Dieter Klingemann. Mapping policy preferences from texts III: statistical solutions for manifesto analysts. Oxford University Press, 2013.

Williams, Christopher, and Jae-Jae Spoon. "Differentiated party response: The effect of Euroskeptic public opinion on party positions." European Union Politics 16.2 (2015): 176-193.

White, Jules, Quchen Fu, Sam Hays, Michael Sandborn, Carlos Olea, Henry Gilbert, Ashraf Elnashar, Jesse Spencer-Smith, and Douglas C. Schmidt. "A prompt pattern catalog to enhance prompt engineering with chatgpt." arXiv 2302.11382 (2023): 1-19.

Wratil, Christopher, and Sara B. Hobolt. "Public Deliberations in The Council of the European Union: Introducing and Validating the DICEU Approach." European Union Politics 20.3 (2019): 511-531.

Ziems, C., W. Held, O. Shaikh, J. Chen, Z. Zhang, and D. Yang. "Can large language models transform computational social science?" Computational Linguistics 50.1 (2024): 237-291.