Who Influences Whom? Measuring the Interplay of Mainstream and Outsider Parties in Social Media Campaigns

Valentina Gonzalez-Rostani*ᴀɴᴅ Claudia Salas Gimenez†

Electoral campaigns are high-frequency environments in which parties receive rapid feedback and face immediate incentives to adjust both message content and presentation. Yet most research on party competition relies on low-frequency texts, such as manifestos, or on election-to-election change, obscuring strategic adaptation within campaigns. We address this gap by analyzing more than 8,000 YouTube transcripts produced by parties in Spain and the United Kingdom during the year before elections between 2015 and 2024. We develop a general framework for studying campaign communication with social media data and recent language models, combining embeddings, BERT-based measures, and large language models. Substantively, we show that parties adapt ideological stances as elections approach, shift issue emphasis, and respond selectively to rivals, especially on particular issues and within ideological families, rather than converging uniformly. Methodologically, our approach captures movement not only in topics, but also in positions, framing, and rhetorical style.

Key words: Social media, YouTube, LLM, embeddings, text-as-data, electoral campaigns, political elites.

*University of Southern California, Department of Political Science and International Relations, gonzalez.rostani@usc.edu
†University of Southern California, Department of Political Science and International Relations, salasgim@usc.edu

INTRODUCTION

Electoral campaigns compress party competition into a short interval of intense strategic adjustment. Parties receive rapid feedback, observe rivals in real time, and revise their messages under severe time pressure. Yet most evidence on party competition comes from low-frequency sources such as manifestos or election-to-election change. These data show that parties respond to voters, competitors, and shifting issue agendas, but they are too coarse to capture adaptation at campaign speed (Klüver and Sagarzazu 2016; Abou-Chadi, Green-Pedersen, and Mortensen 2020; Seeberg 2022; Le Pennec 2024). As a result, we still know surprisingly little about the fine-grained dynamics of campaign interaction: who moves first, who responds, and whether these patterns vary across party families and electoral systems.

This gap is especially consequential in research on mainstream and outsider parties. Existing work shows that competition is shaped by issue entrepreneurship, selective emphasis, and strategic response, but it has largely examined these processes over longer electoral horizons rather than within campaigns themselves (e.g., Meguid 2005).[1] Much of this literature implicitly treats influence as flowing from outsiders to mainstream parties. In studies of immigration, for instance, the central question is often how mainstream parties respond to the radical right (e.g., Gessler and Hunger 2022; Hutter and Kriesi 2022). But this view is incomplete. Other work suggests that the radical right's role in politicizing immigration may at times be overstated, and that mainstream parties themselves often help bring these issues to the center of political conflict (Meyer and Rosenberger 2015; Akkerman 2015). More generally, we know far less about what shapes outsider parties' own campaign behavior, even though radical right and radical left parties have become durable competitors in many party systems. As social media has become a major arena for agenda setting and elite interaction (Hopmann et al. 2012; Barberá et al. 2019; Gilardi et al. 2022; Vesnic-Alujevic and Van Bauwel 2014; Hosseinmardi et al. 2021; González-Rostani, Incio, and Lezama 2025), it is now possible to study campaign influence as a dynamic process and to ask whether it is one-directional, reciprocal, or asymmetrically structured by the competitive environment.

We argue that parties strategically adapt their rhetoric as elections approach, and that their responsiveness to one another is shaped in part by governing viability. Parties with a credible prospect of entering government must appeal to broader electorates, preserve governing credibility, and, in coalition settings, signal compatibility with potential partners. Parties with lower governing viability, such as outsiders, face fewer such constraints and can rely more heavily on sharper rhetoric, ideological differentiation,

---

[1]Examples of work emphasizing longer-run dynamics include Adams and Somer-Topcu (2009), Alonso and Fonseca (2012), Green-Pedersen and Otjes (2019), Ruedin and Morales (2019), Gruber (2014), Heerden et al. (2014), Dalton and McAllister (2015), Meguid (2005), Abou-Chadi and Krause (2020), and Han (2015).

or narrower issue entrepreneurship. This generates an asymmetric competitive structure: office-viable parties are more likely to define the terms of debate, while outsider parties are more likely to adapt their communication in response. But this asymmetry varies across electoral systems. In proportional settings, mainstream parties and some outsiders face coalition-formation constraints (Gonzalez-Rostani 2026) that encourage responsiveness within ideologically proximate blocs. In majoritarian systems, by contrast, mainstream parties compete more directly with one another for swing voters and marginal seats, while outsider parties, which face less credible commitments (Gonzalez-Rostani 2026), maintain distinctiveness through sharper issue differentiation and ideological polarization.

To test these expectations, we turn to YouTube, one of the world's largest and most engaging online media platforms, with 4.95 billion monthly active users (Mohsin 2020; Hosseinmardi et al. 2021). Although research on digital political communication has more often focused on Facebook and X, especially because of their polarizing effects (Conover et al. 2011; Del Vicario et al. 2016; González-Rostani, Incio, and Lezama 2025), YouTube has become an increasingly important source of political information and campaign messaging worldwide (Hosseinmardi et al. 2024; Hosseinmardi et al. 2021; Haroon et al. 2023; Ibrahim et al. 2023; Mamié, Horta Ribeiro, and West 2021). It is especially valuable for studying campaigns because parties use it not only for advertising, but also for mobilization, and narrative construction. We build a new panel of 8,753 videos posted by political parties in Spain and the United Kingdom during the year preceding each general election between 2015 and 2024, using transcripts collected from official party channels through the YouTube Data API. These cases provide a theoretically useful contrast: Spain's proportional system has repeatedly generated coalition bargaining and left-led governing arrangements since 2015, whereas the United Kingdom's majoritarian system produces more direct competition for government among mainstream parties.

We combine several text-analytic tools to measure distinct dimensions of campaign rhetoric. We use BERTopic to estimate issue attention, dictionary-based measures to capture populist discourse, LLM-based scores to recover ideological positioning, and embedding-based measures to estimate rhetorical proximity both to manifesto anchors and to rival parties. This design is important because parties can change not only what they discuss, but also how they frame issues, where they position themselves, and how closely their rhetoric aligns with allies or competitors. Topic models recover issue emphasis, but they do not directly capture stance, ideological signaling, or rhetorical similarity. Embeddings are useful on these dimensions because they place texts in a semantic space that preserves relative meaning. Recent work shows that such representations can recover ideological placement, capture context-dependent meaning, classify political texts across languages, and support comparative analysis (Rheault and Cochrane 2020;

Rodriguez, Spirling, and Stewart 2023; Licht 2023; Widmann and Wich 2023; Wirsching et al. 2025). We therefore treat embeddings not as a stand-alone solution, but as one component of a broader measurement framework that allows us to distinguish shifts in issue attention from shifts in positioning and rhetorical style (Bestvater and Monroe 2023; Linegar, Kocielnik, and Alvarez 2023; Laurer et al. 2024). We then use election-proximity models to study how rhetoric changes as campaigns intensify and daily vector autoregressive models to identify directional interdependence across party families.

Our findings show that parties do adapt as elections approach, but not through simple convergence. In Spain, the strongest late-campaign ideological movement comes from the mainstream left and the radical left, both of which move further left as election day nears, consistent with coalition-oriented adaptation within the bloc with the clearest governing prospects. In the United Kingdom, the largest ideological shift comes from the radical right, which moves further right, consistent with outsider differentiation in a majoritarian setting. Parties also adjust issue emphasis selectively. UK parties increasingly emphasize issues associated with their core electoral identities, whereas Spanish parties shift away from policy discussion toward partisan contestation and coalition framing. The dynamic interdependence results reinforce this interpretation. In Spain, ideological responsiveness is concentrated within the left bloc, where shocks to the mainstream left are followed by movement in the radical left and vice versa. In the United Kingdom, responsiveness is more strongly centered on mainstream parties, and their respond to outsider right.

The paper makes two contributions. Substantively, it brings campaign-time dynamics into the study of party competition by identifying who leads, who follows, as elections draw closer. Parties influence each other, and responsiveness is not reducible to simple outsider-to-mainstream contagion. Methodologically, the paper contributes by developing a framework for studying campaign communication with high-frequency social media data. While work on digital campaigns has relied predominantly on X (e.g., Barberá et al. 2019; González-Rostani, Incio, and Lezama 2025), Facebook (e.g., Poljak and Russell 2025), or lower-frequency sources such as manifestos (e.g., Bräuninger and Giger 2018; Le Pennec 2024), we show how YouTube can be used to study comparative party competition using a broader range of campaign materials. Moreover, we offer a framework that combines topic models, dictionaries, embeddings, and LLM-based scaling to distinguish shifts in issue emphasis from shifts in ideological position and rhetorical style. Because that framework can travel across countries, elections, and platforms, it broadens the study of campaign communication beyond single-issue or single-platform analyses.

The remainder of the paper proceeds as follows. The next section develops the argument about governing viability and asymmetric campaign responsiveness. The following sections describe the

YouTube data, measurement strategy, and empirical design, before presenting the election-proximity and dynamic-interdependence results. The conclusion discusses implications for the study of party competition, digital campaigning, and text-based measurement in comparative politics.

THE DYNAMICS OF STRATEGIC PARTY RESPONSIVENESS

A large literature shows that parties adjust their rhetoric in response to public opinion (Abou-Chadi, Green-Pedersen, and Mortensen 2020), shifts in voter issue priorities (Klüver and Sagarzazu 2016), prior electoral outcomes (Somer-Topcu 2009; Cowburn and Sältzer 2025), and rivals' issue positions (Licht et al. 2025). What remains less clear is which parties lead these adjustments and which parties follow. We argue that the answer turns on governing viability, that is, the extent to which a party can credibly expect to enter government. Governing viability shapes both the constraints parties face in public communication and their capacity to set the campaign agenda. As a result, responsiveness should be asymmetric across party families.

Parties with high governing viability, usually mainstream parties, must appeal to broad electorates, reassure pivotal voters, and protect their reputation for governing competence. In coalition systems, they must also remain acceptable to potential governing partners. These demands reduce the value of sharp rhetorical differentiation and increase incentives for moderation, message discipline, and agenda control. Parties with low governing viability face fewer such constraints and can more readily use sharper or more polarizing rhetoric to differentiate themselves. Yet low-viability parties that aspire to coalition relevance face a tradeoff: they must maintain a distinct electoral profile while avoiding rhetoric that would make future cooperation impossible. Because high-viability parties are more consequential for post-election bargaining and government formation, they are more likely to define the terms of campaign debate; lower-viability parties are more likely to adjust to those moves than to set them independently.

We use this distinction to separate mainstream from outsider parties. Mainstream parties are those that have governed or can plausibly govern and thus possess high governing viability. Outsider parties rely more heavily on differentiation from the mainstream and generally have lower immediate prospects of entering government. This does not mean that outsider parties never shape debate. It does imply, however, that across ideology, populist discourse, and issue attention, mainstream parties should more often lead and outsider parties should more often respond.

This asymmetry is not constant over the electoral cycle. It should intensify as election day approaches. Parties communicate strategically throughout the cycle (Seeberg 2022; Berz and Kroeber 2023), but campaigns compress competition into a short period in which parties receive frequent feedback about

message resonance and rivals' changing agendas. On high-frequency platforms such as social media and online video, this feedback can arrive daily. As the election nears and governing prospects become more immediate, short-run incentives should weigh more heavily than long-run programmatic commitments. Parties should therefore become more willing to shift issue emphasis, adjust populist appeals, and respond directly to rivals' rhetoric. The difference between mainstream leaders and outsider responders should be most visible late in the campaign, when real-time competitive signals are strongest.

Institutional context conditions this dynamic. In proportional, coalition-prone systems, parties compete for votes and for bargaining credibility after the election. Mainstream parties have incentives to signal compatibility with ideologically proximate partners, especially when coalition formation is likely and the election is close (Iversen and Soskice 2006; Sagarzazu and Klüver 2017; Abou-Chadi and Wagner 2019; Bale et al. 2010). As a result, they should move away from manifesto anchors when those anchors conflict with coalition viability. These incentives should be strongest for parties with credible and repeated coalition paths (Gonzalez-Rostani 2026). Outsider parties in these systems still benefit from differentiation, but those with coalition ambitions must also remain responsive to the mainstream parties most likely to lead government. In proportional systems, then, mainstream parties should drive coalition-compatible repositioning, while outsider parties with governing ambitions should respond to mainstream agendas without abandoning their distinct profile.

In majoritarian systems, especially under first-past-the-post, the central contest is over swing voters and marginal seats rather than coalition bargaining (Gonzalez-Rostani 2026). Here mainstream parties should respond primarily to one another, protect their governing viability, and avoid rhetoric that alienates moderate voters. Outsider parties matter mainly through pressure and vote splitting. Their incentives therefore run toward sharper differentiation, more polarizing rhetoric, and emphasis on issues that separate them from mainstream competitors (Van Spanje 2010; Abou-Chadi 2016; Williams and Hunger 2022; Dalton and McAllister 2015). In majoritarian systems, then, mainstream parties should set the terms of debate chiefly in competition with other mainstream parties, whereas outsider parties should seek influence less through agenda leadership than through differentiated appeals that threaten the mainstream's electoral coalition.

This argument complements rather than replaces prominent alternatives. Issue-ownership accounts suggest that parties emphasize issues on which they are seen as competent, while salience-based accounts suggest that parties follow whatever voters currently prioritize (Wagner and Meyer 2014; Klüver and Sagarzazu 2016; Abou-Chadi, Green-Pedersen, and Mortensen 2020). Work on issue engagement further implies that agenda leadership is organizationally structured, with actors closer to party leadership and

government ambition better able to elevate issues (Meyer and Wagner 2021). Likewise, anti-elite rhetoric is not a fixed trait but a strategic resource shaped by electoral returns and office costs (Polk et al. 2017; Licht et al. 2025). Our claim is that these mechanisms are filtered through governing viability. Parties do not simply follow owned issues, public salience, or anti-elite opportunity in the abstract; they do so under constraints that differ systematically between parties likely to govern and parties that are not.

These arguments yield three expectations. First, party responsiveness should be asymmetric: mainstream parties are more likely to be the ones setting campaign agendas, and outsider parties should more often respond to them. Second, parties will adjust their rhetoric as election day approaches. Third, the content of responsiveness should vary across electoral systems: coalition compatibility should matter more in proportional systems, whereas competition for swing voters through differentiation should matter more in majoritarian systems.

RESEARCH DESIGN

This section describes the data, measurement strategy, and empirical design. Figure 1 summarizes the workflow. The analysis proceeds in three stages. First, we construct a dataset of YouTube campaign videos posted to official party channels in Spain and the United Kingdom in the year before each national election between 2015 and 2024, converting video transcripts into observations linked to party, date, and election. Second, we apply a measurement pipeline that captures distinct dimensions of campaign communication: embedding-based cosine similarities (both pairwise and manifesto-anchored), LLM-based ideology scores, dictionary measures of identity and populist discourse, and BERTopic issue shares. These transcript-level measures are aggregated into party-date series. Third, we use these series in two complementary empirical strategies: pooled interaction models that examine whether and how party rhetoric shifts as election day approaches, and daily vector autoregressive models that assess directional interdependence across party families.

Data

To test these hypotheses, we analyze YouTube videos uploaded to official party channels during electoral campaigns, focusing on the year before each election. Spain and the United Kingdom provide a theoretically motivated institutional contrast. Spain's proportional representation system has produced three consecutive left-led coalition governments since 2015, with two elections requiring a second round due to failed coalition formation. The United Kingdom's first-past-the-post system, by contrast, has historically produced single-party majority governments alternating between the Conservative Party on the center-right and the Labour Party on the center-left, generating direct competition for the governing
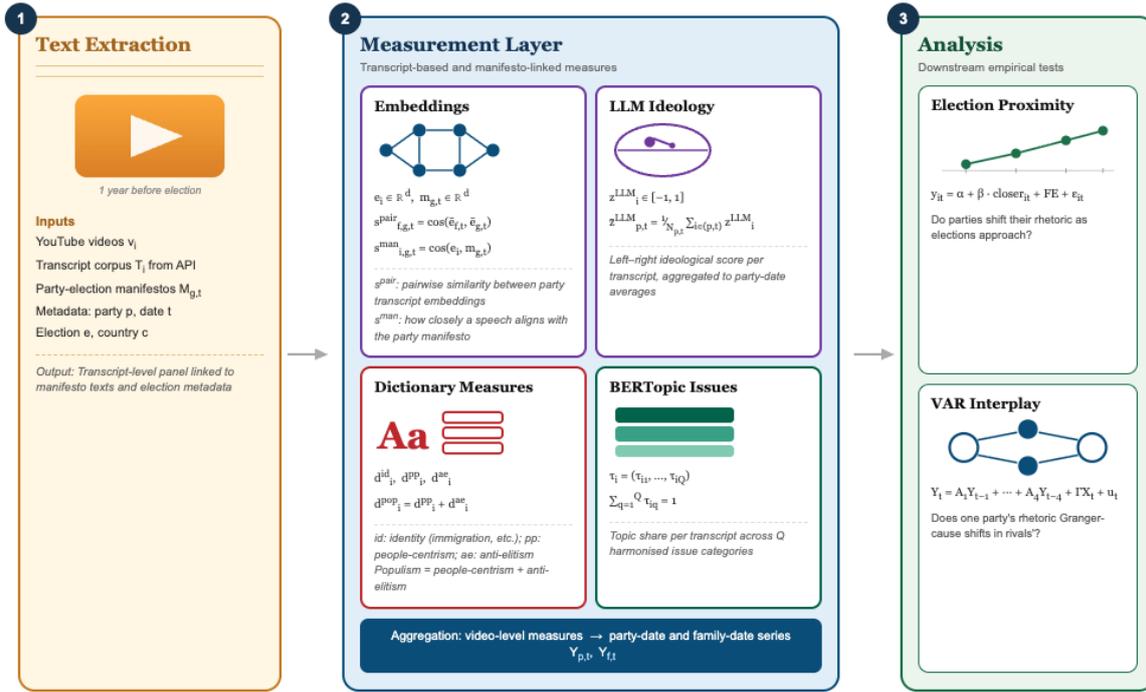
Figure 1: Measurement pipeline.
Note: Party transcripts from YouTube are scored along four dimensions — embedding-based cosine similarities (s^{pair}, s^{man}), LLM ideology (z^{LLM}), dictionary measures of identity and populism (d^{id}, d^{pop}), and BERTopic issue shares (τ_i) — and aggregated to party-date series for election-proximity and VAR analyses.

mandate between two mainstream parties. Together, they allow us to test whether institutional context shapes the direction of asymmetric responsiveness as the theory predicts.

We selected the most competitive parties, grouping them into party families that map onto the distinction between mainstream and outsider parties central to our theoretical framework. As defined above, mainstream parties are those with a realistic prospect of entering government, while outsider parties are those whose electoral appeal rests primarily on differentiation from the mainstream. In our sample, the mainstream left (ML) and mainstream right (MR) families capture the office-viable parties in each country, while the radical left (RL) and radical right (RR) families capture the outsider parties whose governing viability is more limited and whose electoral strategies are more likely to rely on ideological differentiation and issue entrepreneurship.

In Spain, the ML includes PSOE, the MR includes Partido Popular (PP) and Ciudadanos, the RR includes VOX, and the RL includes Podemos, Izquierda Unida, and their coalition Sumar. In the United Kingdom, the ML includes the Labour Party and the Greens,[2] the MR includes the Conservative Party and the Liberal Democrats (center-right leaning), and the RR includes UKIP and Reform UK.

---

[2]Previous scholars have discussed the evolution of the Green party from a niche party focused on the environment to a broader progressive platform, aligning it more closely with mainstream left-wing politics (Dennison 2016).

To analyze the direction and nature of asymmetric responsiveness across time, we draw on a panel of YouTube campaign videos posted to official party channels in the year before each electoral event. The dataset spans the Spanish general elections of 2015, 2016, April and November of 2019, and 2023, and the UK general elections of 2015, 2017, 2019, and 2024, as well as the Brexit referendum of 2016 and the UK local elections of 2021. In total, 8,753 videos were collected across both countries, 6,890 from Spain and 1,863 from the United Kingdom, including regular campaign videos, shorts, and live streams.[3] Table 1 reports descriptive statistics for the sample by party family and country.[4]

TABLE 1: Descriptives YouTube Data by Party Family

Spain					United Kingdom				
Family	Party	Subscribers (K)	Most Popular Video (K)	Total Videos	Family	Party	Subscribers (K)	Most Popular Video (K)	Total Videos
ML	PSOE	64.3	979	1693	ML	Green	32.1	1000	374
MR	PP	120	1500	502	ML	Labour	60	161	100
MR	Ciudadanos	61	3400	988	MR	Conservative	138	3700	307
RL	Izquierda Unida	35.2	269	1333	MR	Liberal Democrats	17.5	357	559
RL	Podemos	168	2400	859	RR	Reform UK	138	523	123
RL	Sumar	11.2	622	95	RR	UKIP	62.7	440	400
RR	VOX	646	4800	1420					
Total				6890	Total				1863

Although YouTube has not been studied as extensively as other social media platforms in the context of electoral campaigns, it offers distinct advantages over the data sources that have dominated existing research. While prior work has relied predominantly on X (e.g., Barberá et al. 2019; González-Rostani, Incio, and Lezama 2025), Facebook (e.g., Poljak and Russell 2025; Bond and Messing 2015), or on low-frequency textual sources such as party manifestos (e.g., Bräuninger and Giger 2018), YouTube is one of the most widely used digital platforms (Hosseinmardi et al. 2021). In Spain, YouTube ranks as the second most popular platform overall, trailing only WhatsApp and ahead of Instagram and Facebook.[5] In the United Kingdom, YouTube is the most widely used online communication platform, reaching 82% of adult internet users, ahead of WhatsApp (80%), Facebook (76%), and Instagram (57%).[6] Beyond its reach, YouTube offers a qualitatively richer source of campaign material than text-based platforms. Party channels host not only short promotional clips but also extended political speeches, rallies, candidate

---

[3]We retained only transcripts containing more than five words to exclude non-coherent text resulting from YouTube's auto-captioning of background noise or music, yielding a final sample of 7,398 usable transcripts out of the 8,753 collected.

[4]For additional descriptive information on the data, including breakdowns by election year, party, and transcript availability, see Appendix Tables A.1.1 through A.1.3.

[5]Observatorio Nacional de Tecnología y Sociedad. (2024). *La Sociedad Digital en España 2024*. Available at https://www.ontsi.es/sites/ontsi/files/2025-01/lasociedaddigital_2024.pdf (accessed March 20, 2026).

[6]The Office of Communications. (2024). *Adults' Media Use and Attitudes Report 2024*. Available at https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/adults/adults-media-use-and-attitudes-2024/adults-media-use-and-attitudes-report-2024.pdf (accessed March 20, 2026).

interviews, debate reactions, and live streams. Prior research has shown that parties use YouTube not only for advertising but also to mobilize supporters and communicate policy positions (e.g., Vesnic-Alujevic and Van Bauwel 2014; Sohal and Kaur 2018; Scherr, Reinemann, and Jandura 2015), and more recent work has begun to leverage the platform's transcript content to estimate ideological positioning (Lai et al. 2024). We build on this emerging line of research by applying a measurement pipeline to capture issue attention, rhetorical strategy, ideology, and identity and populist framing, offering a more complete account of the multiple dimensions of party communication during campaigns.

Data collection proceeded in two steps. First, we scraped official party channels to identify all videos posted during the period of analysis and retrieve their unique video IDs. Second, we used the YouTube Data API to collect the available auto-generated transcripts for each identified video.[7] We successfully obtained transcripts for 8,753 videos (84% of all videos collected), 82% of the videos from Spain and 93.6% from the United Kingdom, which constitute the final sample.[8] To give a sense of the corpus size, the full transcript sample amounts to approximately 855,325 tweet-length segments (140 characters each), reflecting the substantially greater volume of textual content that YouTube transcripts provide relative to the short-form posts typical of X-based studies.

From Text to Rhetorical Strategies

We use a set of complementary text measures to capture different key dimensions of party communication in campaign videos. No single indicator is sufficient for this task: parties may discuss the same issue with different rhetorical frames, signal similar ideological positions through different vocabularies, or shift their emphasis across topics while maintaining a stable programmatic profile. Relying on any one measure alone risks missing important aspects of strategic adaptation. To address this, we combine embedding-based cosine similarity, LLM-based ideology scores, and topic measures derived from BERTopic and dictionaries. Together, these measures provide a broader account of campaign rhetoric, allowing us to study not only what parties talk about, but also how they position themselves relative to rivals and their own longer-term programmatic commitments.

Rhetorical Proximity (Embedding-Based Similarity). To capture how closely parties' campaign rhetoric aligns with that of rivals and with their own longer-run programmatic commitments, we represent each transcript as a dense vector in a shared semantic space and summarize relationships between texts using cosine similarity. The underlying logic is distributional: words and phrases that appear in similar

---

[7]See the Appendix A.2 for further details on the data collection process.

[8]Transcript availability is distributed evenly across parties within each country. Further details can be found in Appendix Table A.1.3.

contextual environments are placed closer together in vector space, so proximity between two texts reflects semantic similarity rather than exact word overlap. This is particularly well suited to campaign analysis, where parties may discuss the same issue through different rhetorical frames, signal similar ideological positions through different vocabularies, or shift their rhetoric in substantively meaningful ways without changing their topic labels. Embeddings therefore allow us to compare campaign videos at the level of contextual meaning rather than surface-level term frequencies.

We construct two quantities from these representations, both illustrated in Figure 2. The first is pairwise rhetorical similarity, defined as the cosine similarity between the transcript embeddings of two party families, which captures how semantically proximate their campaign communication is at a given point in time. The second is manifesto congruence, defined as the cosine similarity between a party's transcript embedding and the embedding of its party-family manifesto, which captures how closely campaign rhetoric aligns with longer-run programmatic anchors. Together, these two quantities allow us to distinguish between convergence toward rivals and divergence from a party's own stated program, both of which are theoretically meaningful forms of strategic adaptation.

(a) Video-to-video similarity
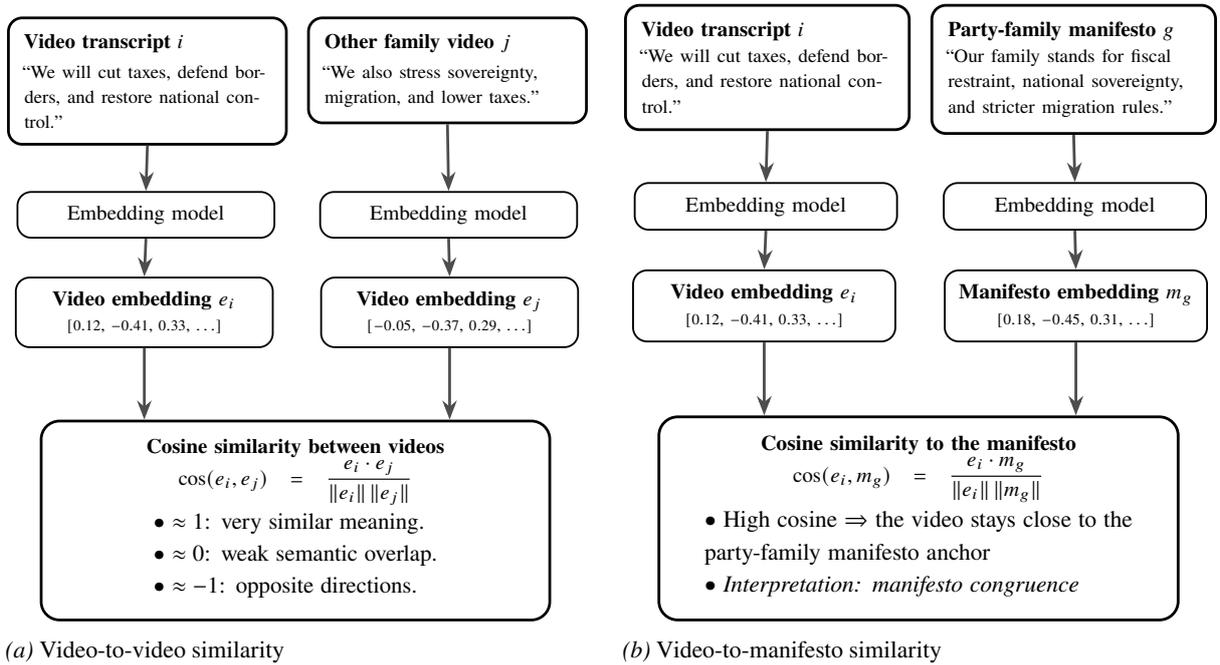
(b) Video-to-manifesto similarity

Figure 2: How embedding-based similarity is constructed from campaign texts. In the left panel, two video texts are embedded and compared using cosine similarity. In the right panel, a video embedding is compared with the manifesto embedding of a party family.

Formally, let e_i denote the embedding of transcript i, e_j the embedding of another transcript, and m_g the embedding of the manifesto associated with party family g. To generate these embeddings, we encode each transcript using a pretrained multilingual sentence transformer model, which maps variable-length

texts to fixed-dimensional dense vectors in a shared semantic space. Pairwise rhetorical similarity is then computed as the cosine of the angle between two transcript vectors, $\cos(e_i, e_j) = \frac{e_i \cdot e_j}{\|e_i\|\|e_j\|}$, and manifesto congruence as the cosine similarity between a transcript vector and the corresponding manifesto embedding, $\cos(e_i, m_g) = \frac{e_i \cdot m_g}{\|e_i\|\|m_g\|}$. Higher values indicate greater semantic alignment, values near zero indicate weak overlap in meaning, and lower values indicate more divergent rhetorical content. These scores are the quantities used to track movement toward rivals, coalition-relevant actors, or manifesto anchors over the course of the campaigns. Figure 3 illustrates both measures for the 2023 Spanish general election.
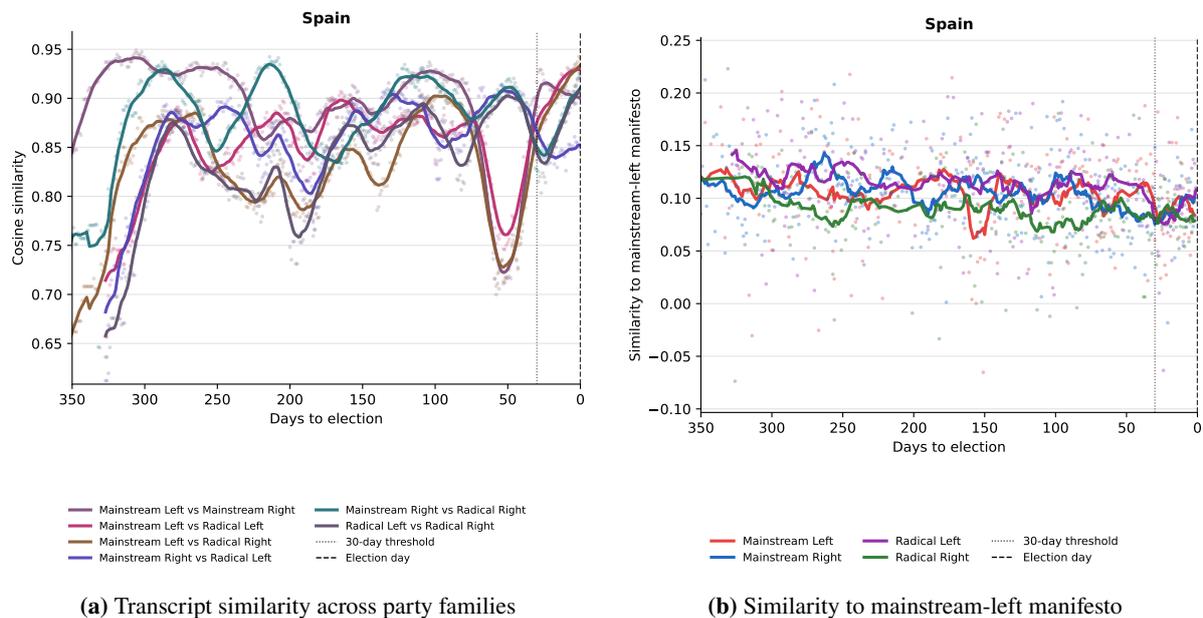

(a) Transcript similarity across party families

(b) Similarity to mainstream-left manifesto

Figure 3: Rhetorical similarity in Spain during the final year before the July 23, 2023 general election.

Notes: Panel (a) reports cosine similarity between 30-day trailing transcript-embedding centroids aggregated to the party-family level. The six pairs shown are Mainstream Left vs Mainstream Right, Mainstream Left vs Radical Left, Mainstream Left vs Radical Right, Mainstream Right vs Radical Left, Mainstream Right vs Radical Right, and Radical Left vs Radical Right. Panel (b) reports daily transcript similarity to the mainstream-left manifesto by party family. In both panels, faint points show daily observations and solid lines show smoothed trends (15-day rolling averages in panel (a), 30-day in panel (b)). The dotted vertical line marks 30 days before the election, and the dashed vertical line marks election day.

This approach builds on recent methodological work showing that embeddings can recover latent political concepts, place actors in a common semantic space, and support inference about substantively meaningful semantic differences across groups and periods (Rheault and Cochrane 2020; Rodriguez, Spirling, and Stewart 2023). Embeddings also provide compact features for downstream analysis, permit qualitative validation through nearest-neighbor and concept-comparison checks, and can rely on pretrained multilingual models when local corpora are limited.

Ideological Positioning (LLM Ideology Scoring). We estimate ideological positioning by directly prompting a large language model to locate each campaign transcript on a common left–right scale.

The key rationale behind this approach is that ideological signals in campaign communication are often implicit, context-dependent, and expressed through varied language rather than fixed keywords. Methods based solely on dictionaries or supervised classification can struggle in this setting, particularly when texts differ in length, style, or language, or when comparable labeled training data are not available across countries and periods (Mens and Gallego 2025). By contrast, direct prompting allows us to recover ideology as an interpretable, continuous dimension without requiring a task-specific training set, while maintaining comparability across heterogeneous corpora.[9]

We implement this using OpenAI's `gpt-4o-mini` model via the Responses API, prompting the model to assign each transcript a position on a 0–100 scale, where 0 denotes the extreme left, 50 the center, and 100 the extreme right; when a transcript does not contain meaningful ideological content, the model returns a missing value.[10] We set the temperature to zero and restrict outputs to numeric responses in order to maximize consistency and reproducibility across the full corpus of transcripts. This produces a transcript-level ideology score, z_i^{LLM}, which we aggregate to the party-date level for the time-series analyses below. Figure 4 shows average transcript-based ideology scores by party family and country. The measure recovers expected patterns: radical right scores highest, the radical left lowest, and mainstream parties fall in between; in the United Kingdom, where no radical left party is in the sample, the radical right scores highest, followed by the mainstream right and then the mainstream left. Additional descriptives appear in Appendix B.2.
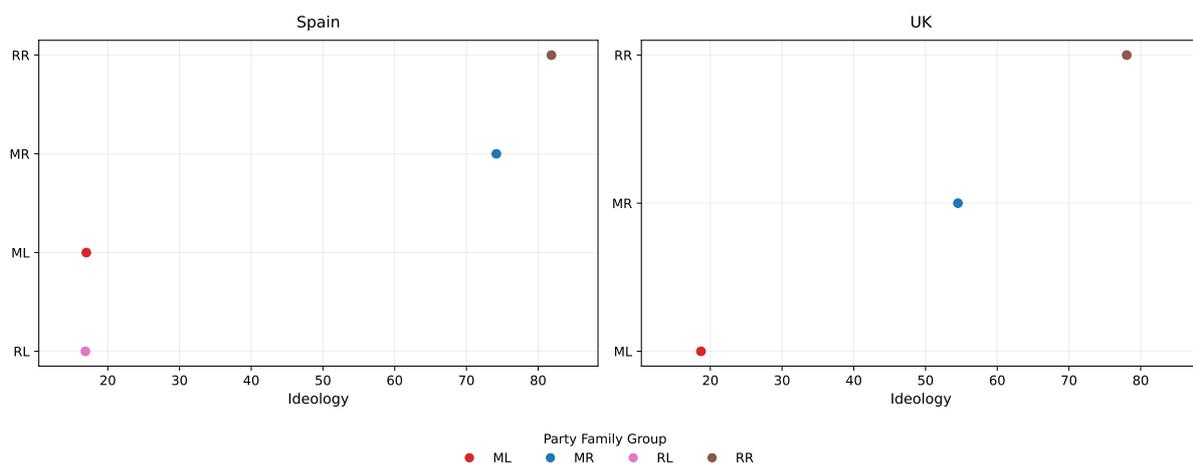

Figure 4: Average Levels of Ideology by Party Family and Country

Notes: Points indicate the count average of transcript-based AI ideology scores for each party family within each country. Ideology scores range from 0 (extreme left) to 100 (extreme right). Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right.

---

[9]Recent work shows that instruction-tuned language models recover policy positions and topics from political texts (Mens and Gallego 2025; González-Rostani, Incio, and Lezama 2025).

[10]The exact wording and settings of the prompt are reported in Appendix B.1.

Issue Attention using BERTopic. To measure issue attention, we use BERTopic, an unsupervised topic-modeling approach that combines transformer-based embeddings, clustering, and class-based TF–IDF to recover coherent themes from large corpora (Grootendorst 2022). We use the resulting assignments to construct a transcript-level topic measure, τ_i, that captures the relative emphasis each party places on the discovered issues.[11] Because it relies on contextual embeddings rather than bag-of-words counts, BERTopic is well suited to campaign transcripts, where issue agendas are often expressed through short, context-dependent language. Recent work shows that these models outperform classical approaches that do not leverage transfer learning (Laurer et al. 2024), and their application to political text is a recent but rapidly expanding area of research (Bestvater and Monroe 2023; Licht 2023; Widmann and Wich 2023; González-Rostani, Incio, and Lezama 2025).

Identity and Populism with Dictionaries. We complement BERTopic with dictionary-based measures for constructs that are theoretically narrower and more interpretable than broad issue agendas. Dictionary methods remain valuable when the research question concerns explicit rhetorical content that can be defined ex ante, such as references to collective identity, pro-people language, or anti-elite attacks. In this sense, dictionaries provide a useful theory-driven counterweight to more flexible transformer-based approaches. We rely on previous scholars' dictionary-based approaches (Rooduijn and Pauwels 2011; Gonzalez-Rostani 2026).[12] We score each transcript with dictionaries capturing identity-related language and the main rhetorical components of populist discourse, especially pro-people and anti-elite appeals. These transcript-level scores, denoted d_i^{id} and related category-specific dictionary measures, are based on the presence of terms and expressions associated with each category and are then aggregated to the party-date level.[13]

Empirical Analysis

We employ two complementary empirical strategies to study how parties strategically adapt their campaign communication and respond to one another over the course of the campaign. The first examines behavioral adjustment over the campaign timeline, asking whether parties systematically recalibrate their ideological positioning, issue emphasis, and rhetorical framing as election day approaches. The second examines directional interdependence across party families, asking whether observed co-movement in party rhetoric reflects symmetric responses to common shocks or asymmetric responsiveness in which some party families consistently lead and others follow. Together, these two strategies allow us to distinguish

---

[11]Descriptive figures for the distribution of BERTopic issue shares by party family and country are reported in Appendix C.1.

[12]Appendix D reports the full list of words and expressions included in the dictionaries.

[13]Appendix Figures D.4.1 and D.4.2 show the distribution of populist and identity rhetoric across party families in each country.

within-party adaptation from cross-party influence and to evaluate our theoretical expectations about the role of governing viability and institutional context in shaping campaign dynamics.

Evaluating Rhetorical Shifts Near Election Day. To study how party communication evolves as elections approach, we restrict the analysis to country-specific election windows spanning the year before each electoral event through election day. The unit of observation is the party-date. Within each window, we construct a continuous measure of electoral proximity defined as the number of calendar days between the date on which a video was posted and election day, rescaled so that higher values indicate greater proximity to the election.

We estimate pooled linear models that relate each outcome of interest to electoral proximity, allowing for heterogeneous responses across party families. The specification includes party-family indicators and their interactions with the proximity measure, capturing both baseline differences across families and differences in how they adjust their communication as the election nears. Election-period fixed effects account for common shocks within each campaign, and a country indicator for the United Kingdom absorbs cross-country differences in baseline levels. The estimating equation is:

$$y_{im} = \alpha_m + \beta_m \cdot closer_i + \sum_{f \neq f_0} \gamma_{fm} D_{if} + \sum_{f \neq f_0} \delta_{fm}(closer_i \cdot D_{if}) + \sum_{e \neq e_0} \eta_{em} P_{ie} + \theta_m \cdot UK_i + \varepsilon_{im}$$

where the coefficient β_m captures the average rate of change in outcome m as the election approaches for the reference party family, and the interaction terms δ_{fm} recover family-specific deviations from this baseline trend.

Evaluating Who Sets the Terms of Debate. While the election-proximity models capture how individual party families adjust their rhetoric over time, they do not directly address who responds to whom. Party communication is inherently interdependent: changes in one party's messaging may prompt responses from rivals, and those responses may in turn feed back into the original party's communication. Shifts in issue attention or rhetorical appeals are also likely to exhibit persistence, as past emphasis shapes current communication choices. Studying these dynamics requires an empirical framework that can accommodate multiple endogenous variables evolving jointly over time and allow the data to reveal the direction of influence rather than imposing it a priori.

To address this, we estimate vector autoregressive (VAR) models with electoral-campaign fixed effects.[14] VAR models allow each variable to depend on its own past realizations as well as on the lagged values of all other variables in the system, providing a flexible way to capture strategic interdependence

---

[14]Electoral-campaign fixed effects account for period-specific factors common to all parties within a given campaign.

without imposing strong assumptions about the direction of influence (Freeman, Williams, and Lin 1989). This approach has been widely used to study dynamic interactions among political actors and communication processes (Barberá et al. 2019; Gilardi et al. 2021; Brandt, Colaresi, and Freeman 2008; Brandt and Freeman 2009; Edwards and Wood 1999; Enders and Sandler 1993). Unlike structural models that require specifying a causal ordering in advance, VAR treats all included variables symmetrically and lets the temporal structure of the data reveal their relationships (Qin 2011).

We model a system of stationary daily time series capturing party-level communication across the dimensions described above. Let $Y_{i,j,t}$ denote the level of attention that party family i devotes to dimension j on day t. Because these measures are skewed–with occasional spikes and many low-attention observations–we apply a logarithmic transformation, defining $Z_{i,j,t} = \log(Y_{i,j,t} + 1)$. The transformed series for each party family evolves as a function of its own past values and the lagged values of all other party families in the system, capturing both persistence in communication and cross-party dynamics. We estimate:

$$Z_{i,j,t} = \alpha_e + \sum_k \sum_{p=1}^4 \beta_{k,p} Z_{k,j,t-p} + \varepsilon_{i,j,t},$$

where α_e denotes electoral-campaign fixed effects and the outer sum runs over all party families k in the system. The lag length of four days is selected based on Akaike Information Criterion (AIC) minimization, and results are robust to alternative lag structures. To interpret the estimated relationships, we compute cumulative impulse response functions (IRFs), which trace how a one-standard-deviation shock to one party family's communication propagates across all other party families over a 15-day horizon, allowing us to assess both the direction and the magnitude of asymmetric responsiveness in campaign rhetoric.

Results

We organize the results around two questions: how party rhetoric changes as election day approaches, and which party families' shifts are followed by responses from others. Across measures, campaign rhetoric becomes more strategically responsive late in the campaign, but not through uniform convergence on a common rhetorical position. Instead, the content and direction of adjustment vary with party family, government-formation prospects, and institutional context. The United Kingdom exhibits a competitive structure centered on mainstream parties, whereas Spain shows stronger within-bloc adjustment, especially on the left.

Strategic Rhetorical Shifts Near Election Day

We first examine whether campaign rhetoric changes systematically in the run-up to the election. Figure 5 plots the similarity between campaign transcripts and manifesto language for each party family. The common pattern is declining transcript-to-manifesto similarity, which suggests that campaign discourse becomes less anchored in pre-campaign programmatic commitments as election day nears. This decline is strongest in Spain, where the radical right, and to a lesser extent the radical left, move furthest from their manifesto baselines. In the United Kingdom, the pattern is weaker and less uniform, although the radical right again departs clearly from manifesto language.
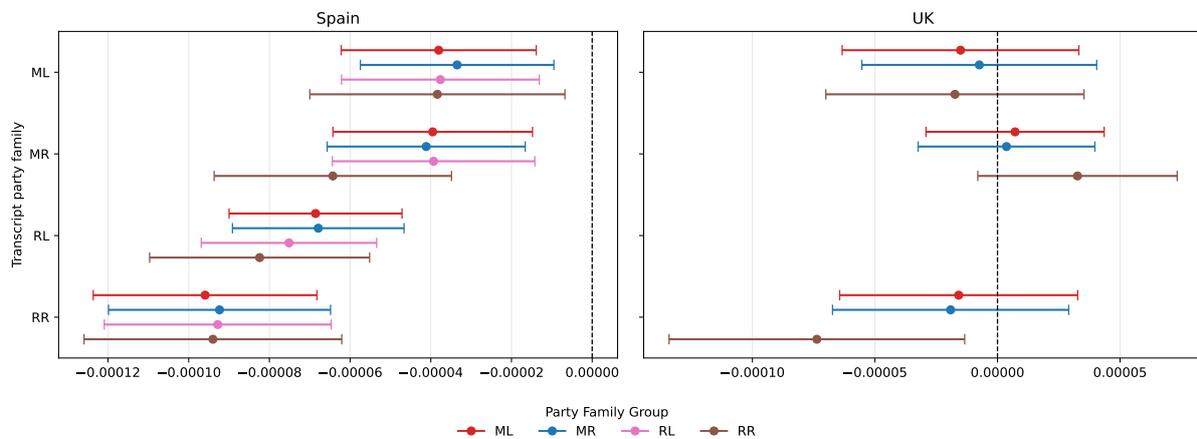
Figure 5: Election Proximity and Transcript-to-Manifesto Similarity by Party Family

Notes: The dependent variable is transcript-to-manifesto cosine similarity, with larger values indicating greater similarity. Party families are shown on the y-axis, and manifesto target families in the legend. Each point reports the estimated effect of one additional day closer to the election from pooled interaction models with party-family fixed effects, election-period fixed effects, and a country fixed effect for the United Kingdom (Spain is the reference category). Larger values of the independent variable indicate greater proximity to election day, so more negative coefficients imply lower transcript-to-manifesto similarity as elections approach. The estimation sample contains $N = 4,008$ observations for ML and MR targets, $N = 3,236$ for RL targets, and $N = 2,670$ for RR targets.

A decline relative to manifesto language does not reveal whether parties move toward one another. Figure 6 addresses this question by examining pairwise transcript similarity across party families. Here the dominant pattern runs in the opposite direction: party families generally become more similar to one another as election day approaches. This tendency is strongest in the United Kingdom, where similarity increases across all party-family pairs. Spain shows the same general tendency, but with smaller and less consistent effects. Taken together, the first two figures suggest that late-campaign rhetoric becomes less anchored in parties' own manifestos and more responsive to the evolving competitive environment.

The Substantive Direction of Late-Campaign Change   These semantic shifts do not by themselves identify the substantive direction of movement. Greater pairwise similarity need not imply convergence
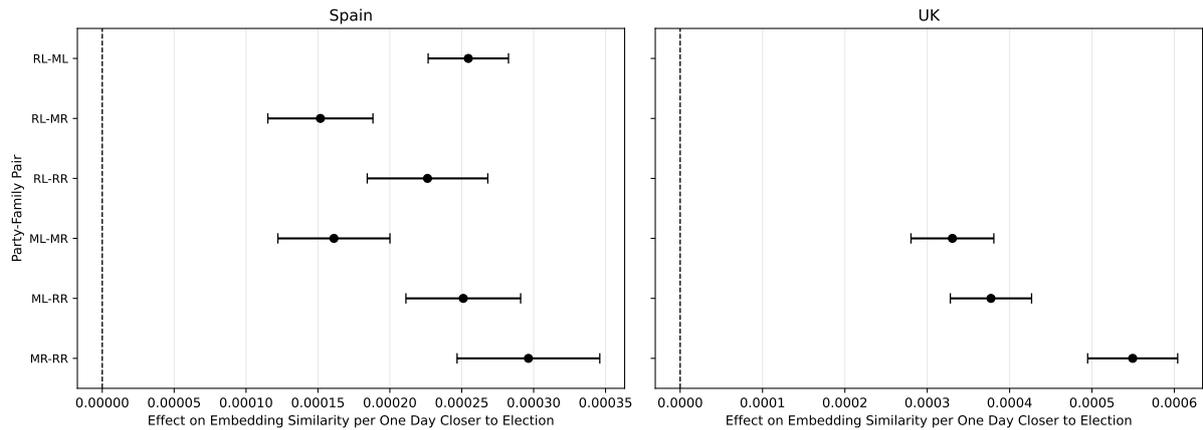
Figure 6: Embedding Similarity by Election Proximity and Country

Notes: The dependent variable is transcript cosine similarity for a given party-family pair (for example, ML–RR), estimated separately by country and pair, with larger values indicating greater similarity. The main independent variable is days closer to election, defined so that larger values indicate observations closer to election day. Models include election-period fixed effects. The pair-specific estimation sample ranges from $N = 1,384$ to $1,477$ in Spain and from $N = 1,463$ to $1,942$ in the United Kingdom.

on ideology, populist style, or issue priorities. We therefore turn to those dimensions directly. Figure 7 examines ideological positioning. The results do not show convergence on a common left-right location. Instead, ideological movement is structured by party family and country. In Spain, both the mainstream left and the radical left shift further left as election day approaches. In the United Kingdom, the clearest movement comes from the radical right, which shifts further right. The increase in pairwise similarity documented above, therefore, does not reflect general ideological moderation. Rather, it is more consistent with shared attention to the campaign environment than with movement toward the ideological center.

The cross-national contrast is also substantively informative. In Spain, ideological adjustment is concentrated within the left bloc, the coalition space with the strongest governing potential during the periods we study. In the United Kingdom, late-campaign ideological change is driven mainly by the radical right, while mainstream parties remain comparatively stable. Where coalition bargaining is salient, adjustment appears to occur within a plausible governing bloc. Where outsider parties lack realistic governing prospects, they preserve, and in some cases sharpen, ideological distinctiveness.

Turning to specific rhetorical cues, Figure 8 points to the same conclusion: late-campaign adjustment is differentiated rather than uniform. In Spain, change is concentrated on the left. The outsider left increases identity and pro-people rhetoric while reducing anti-elite cues. On the right, most populist cues decline as election day approaches, with identity rhetoric as the main exception, consistent with the radical right's stable emphasis on immigration and related themes. In the United Kingdom, the radical right increases identity rhetoric, whereas both mainstream parties reduce populist appeals across most
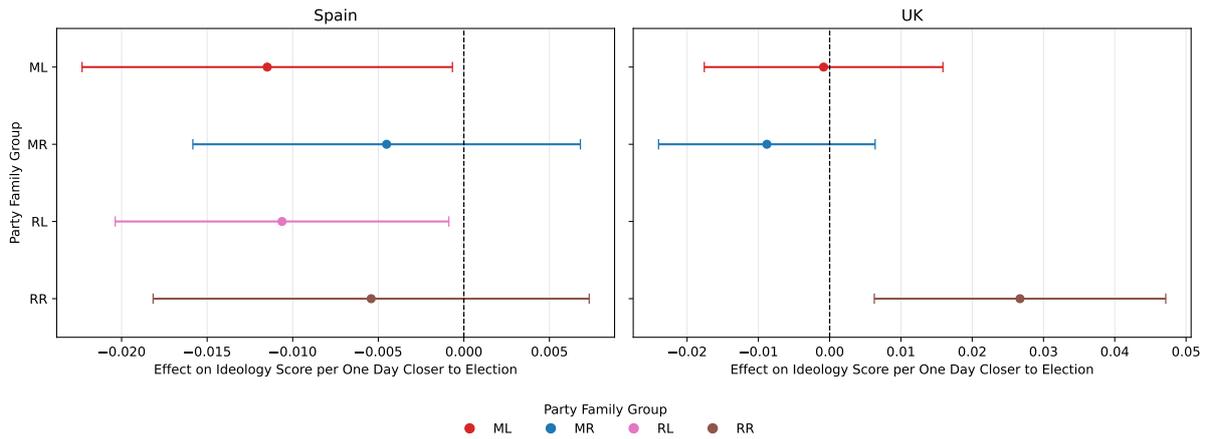
Figure 7: AI Ideology by Election Proximity, Party Family, and Country

Notes: The dependent variable is the transcript-based AI-based ideology score, aggregated as a daily mean by date × country × party. The score ranges from the extreme left (0) to the extreme right (100). The main independent variable is days closer to the election, with larger values indicating greater proximity to election day. Estimates come from a pooled interaction model with party-family fixed effects, election-period fixed effects, and a country fixed effect for the United Kingdom (Spain is the reference category). The plotted coefficients are country-family-specific slopes recovered from interactions between election proximity and party family. The estimation sample contains $N = 4,044$ observations. The unit of observation is party-date.

measures. These results are consistent with our expectations that campaign adaptation is structured by party incentives: mainstream parties moderate potentially costly appeals, while outsider parties retain sharper rhetorical profiles.
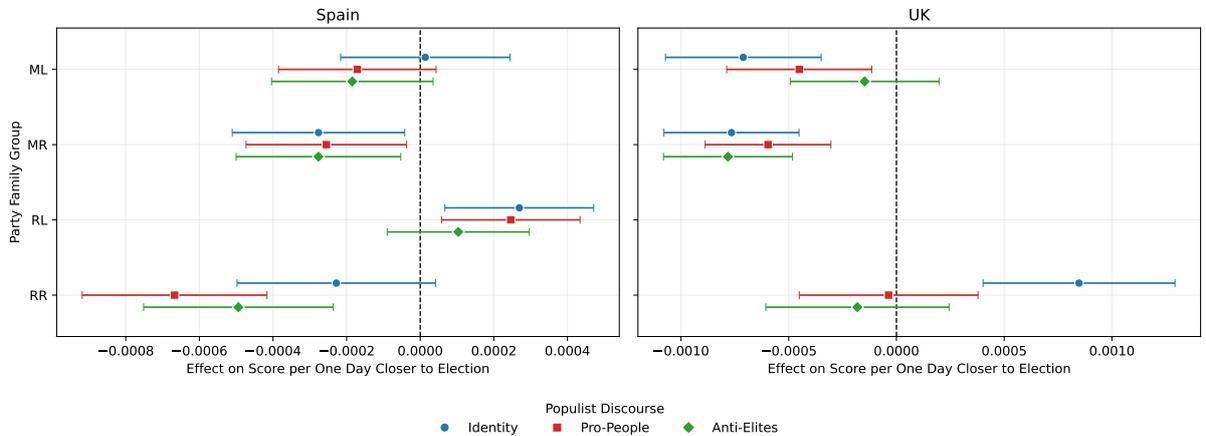

Figure 8: Dictionary-Based Populist Discourse by Election Proximity, Measure, Party Family, and Country

Notes: The dependent variables are transcript-based dictionary measures of identity, pro-people rhetoric, and anti-elite rhetoric. The main independent variable is a continuous measure of days closer to election, constructed as the number of calendar days between the video posting date and rescaled so that larger values indicate greater proximity to election day. Estimates come from pooled interaction models with party-family fixed effects, election-period fixed effects, and a country fixed effect for the United Kingdom (Spain is the reference category). The broader specification also includes additional dictionary outcomes, but this figure reports only the three measures shown. The estimation sample contains $N = 4,821$ observations. The unit of observation is party-date.

Figure 9 turns to issue attention. In the United Kingdom, party families increasingly emphasize issues they are comparatively well positioned to own as election day approaches. The mainstream left shifts

toward the environment and housing, the mainstream right toward taxes and fiscal policy, and the radical right toward immigration. In Spain, by contrast, late-campaign adjustment is less about stable issue ownership than about conflict over government formation. As election day nears, attention shifts toward interparty conflict, coalition bargaining, and ideologically charged attacks, while emphasis on issues such as the economy or defense declines. Mainstream parties increasingly frame the election in terms of governing stakes and stability; the radical left emphasizes agreements among left parties; and the radical right directs attention toward criticism of the governing coalition.[15] The contrast across countries is clear. In the United Kingdom, late-campaign differentiation operates mainly through issue emphasis. In Spain, adaptation is more closely tied to coalition signaling and partisan contestation over who can govern.

Who Sets the Terms of Debate?

The close-to-election analyses identify systematic change over the campaign, but they do not reveal directional dependence. We therefore turn to VAR models, which ask whether changes in one party family's rhetoric are followed by changes in others over subsequent days.

Figure 10 reports the VAR results for ideological positioning. In the United Kingdom, responses are limited and concentrated around the mainstream. The mainstream left responds to shocks from the radical right, but the radical right does not respond to shifts by other party families. In Spain, by contrast, ideological shocks on the left generate clear within-bloc responses: movements by the mainstream left are followed by movements by the radical left, and vice versa. Dynamic responsiveness is therefore strongest within the left bloc, again consistent with the coalition structure that made this bloc the most viable governing alternative during the period under study.

Figure 11 shows a related, though broader, pattern for populist discourse.[16] In the United Kingdom, responsiveness is again organized largely around competition involving the mainstream parties. The two mainstream parties respond to one another across most rhetorical dimensions and, in some instances, to shocks from the radical right. For example, increases in radical-right identity rhetoric are followed by reductions in identity rhetoric by the mainstream left. In Spain, responsiveness is more cross-cutting. The mainstream left responds to shocks from the radical right across populist dimensions, while the

---

[15]For instance, mainstream parties on both sides increasingly frame the election in terms of governing stakes and partisan conflict (e.g., "Next Sunday, if you want stability, if you want a strong government, and if you want a progressive government that stops the extreme right, the Socialist Party is your home," PSOE, 8 November 2019; "We are at risk that there will be Chivite nationalists and communists or that there will be a government of the Popular Party that fights," PP, 8 November 2019). Parties on the left outside the mainstream emphasize coalition formation and agreements (e.g., "Today I am pleased that we have been able to reach an agreement with the Socialist Party, with the United Left, and with Compromise...," Podemos, 15 March 2016), while parties on the outsider right focus on criticizing the governing coalition (e.g., "Spain needs a serious and decent social democratic party because Spain doesn't have one," VOX, 21 March 2023).

[16]VAR results for BERTopic harmonized topic shares are reported in Appendix Figure F.1.1.
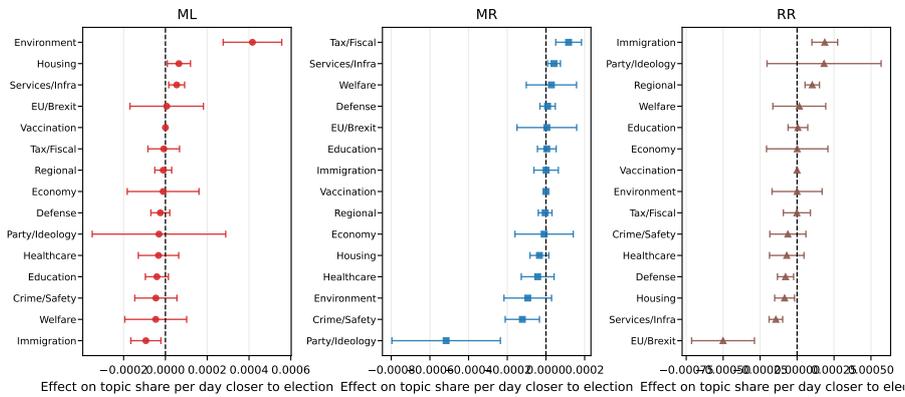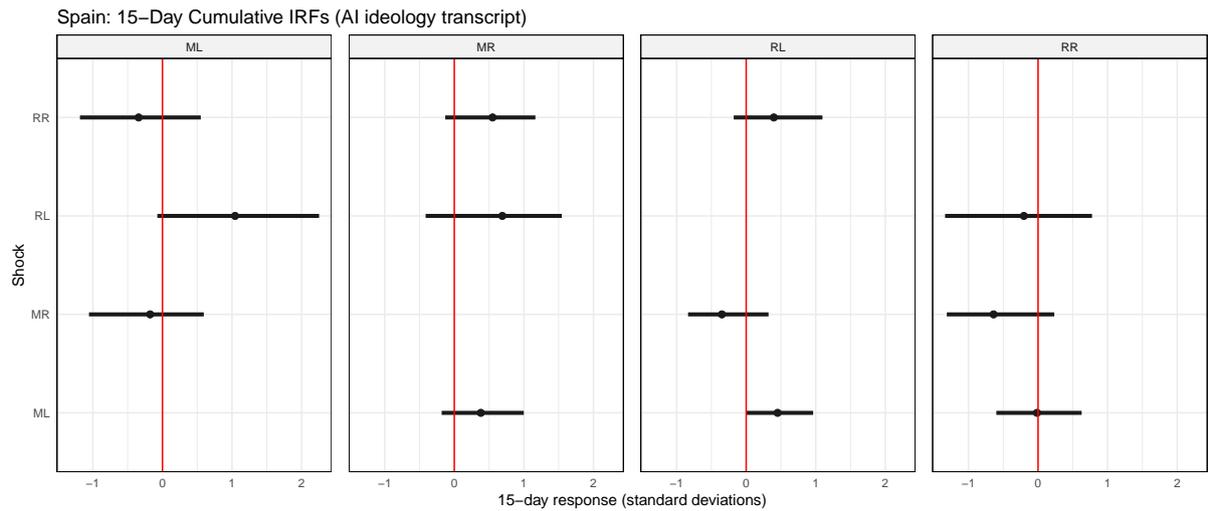
(a) Spain
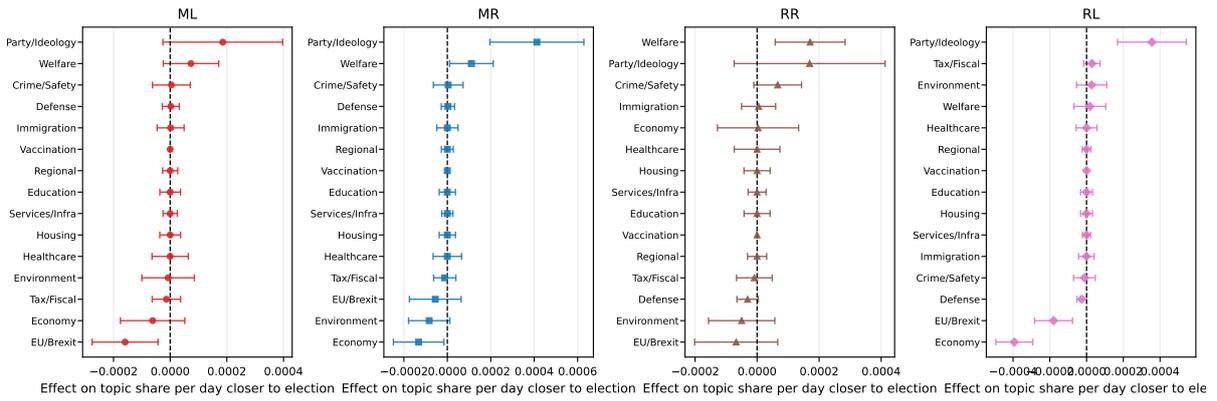
(b) United Kingdom

Figure 9: BERTopic and Election Proximity by Country, Party Family, and Issue
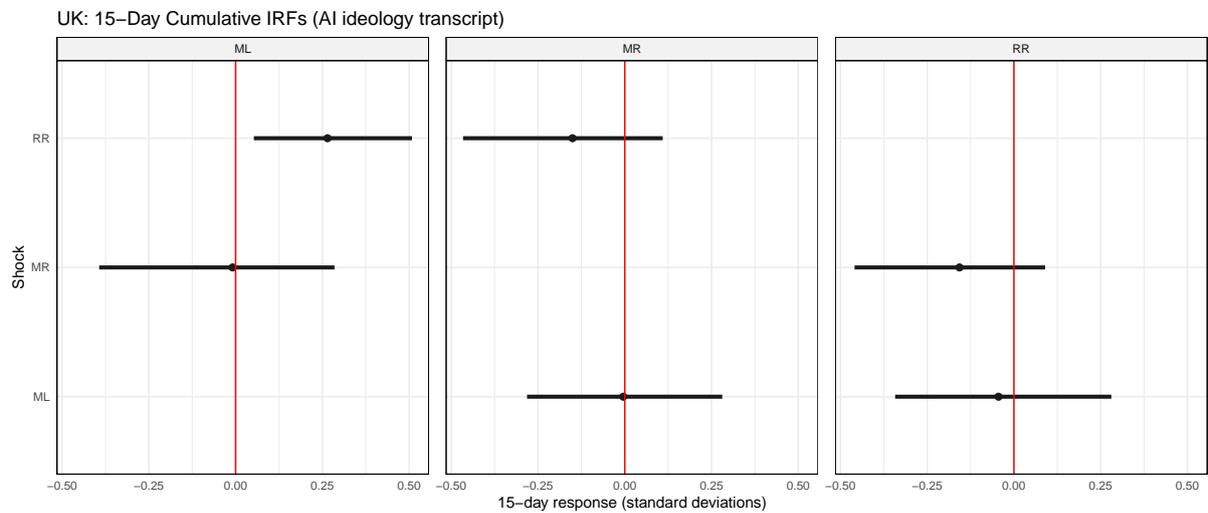
Notes: The dependent variables are transcript-based BERTopic harmonized topic shares. The main independent variable is a continuous measure of days closer to election, constructed as the number of calendar days between the video posting date and election day and rescaled so that larger values indicate greater proximity to election day. Estimates come from pooled interaction models with party-family fixed effects, election-period fixed effects, and a country fixed effect for the United Kingdom (Spain is the reference category). The estimation sample contains $N = 4{,}276$ daily family observations. Within each country-family panel, topics are ranked from the largest to the smallest estimated close-election effect.

mainstream right responds to the radical right mainly on identity rhetoric. Among outsider parties, the radical left responds to the radical right, and the radical right, especially through anti-elite rhetoric, responds to shocks from the mainstream left.

Overall, the results indicate that late-campaign rhetoric is strategic, but its form varies across party systems and issues. In the United Kingdom, adjustment is organized primarily around competition involving the mainstream parties. In Spain, adjustment is more relational and more sensitive to coalition geometry, especially within the left bloc. Across both settings, outsider parties are not simply agenda setters operating on their own terms; they also adapt to strategically important rivals, even when they maintain distinct ideological or issue-based profiles.

Spain: 15–Day Cumulative IRFs (AI ideology transcript)

(a) Spain

UK: 15–Day Cumulative IRFs (AI ideology transcript)

(b) United Kingdom

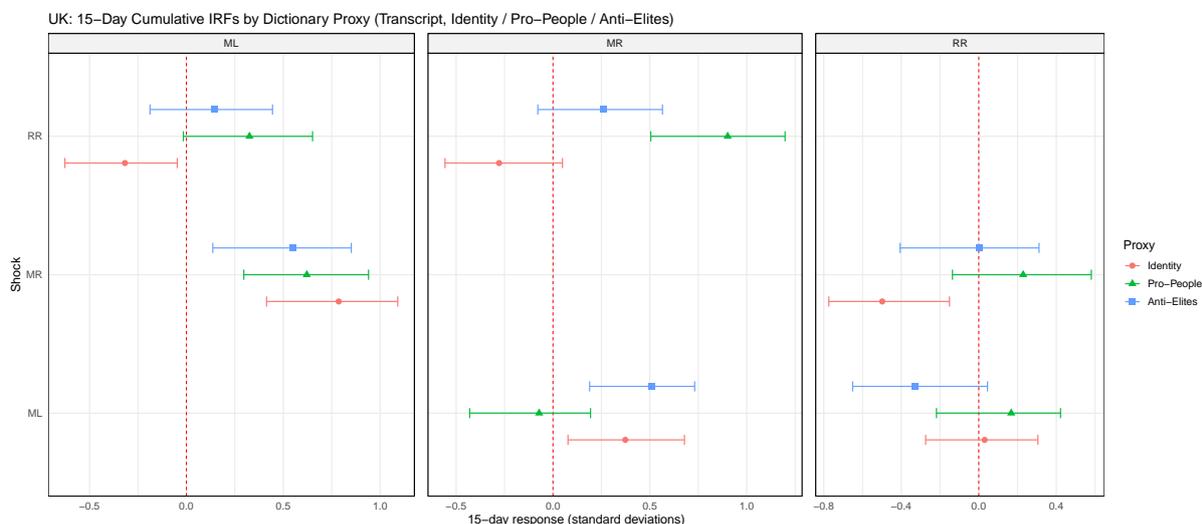*Figure 10:* VAR Results for AI Ideology in Spain and the United Kingdom

Notes: Panel (a) reports Spain and panel (b) the United Kingdom. The dependent variables are daily AI ideology series constructed from YouTube transcripts, aggregated as a daily mean by party family, and ranging from 0 (extreme left) to 100 (extreme right). The figures display cross-group 15-day cumulative impulse-response estimates from separate VAR(4) models, excluding own-response terms. The underlying variables are daily AI ideology series constructed from YouTube transcripts, ranging from 0 (extreme left) to 100 (extreme right). In each specification, each series enters as both a dependent variable and a lagged predictor, and election-period dummies are included. Points denote standardized effects and horizontal bars indicate bootstrap confidence intervals. The estimation sample includes $N_{\text{Spain}} = 1,483$ daily observations and $N_{\text{UK}} = 2,182$ daily observations.

FINAL REMARKS

Electoral campaigns compress party competition into a short interval of intense strategic adjustment. Yet most evidence on party competition still comes from low-frequency sources that are not well suited to capturing adaptation at campaign speed. This paper addressed that gap by examining who influences whom during campaigns, whether responsiveness is reciprocal or asymmetric, and how those dynamics vary across party families and electoral systems. Using YouTube communication from Spain and the
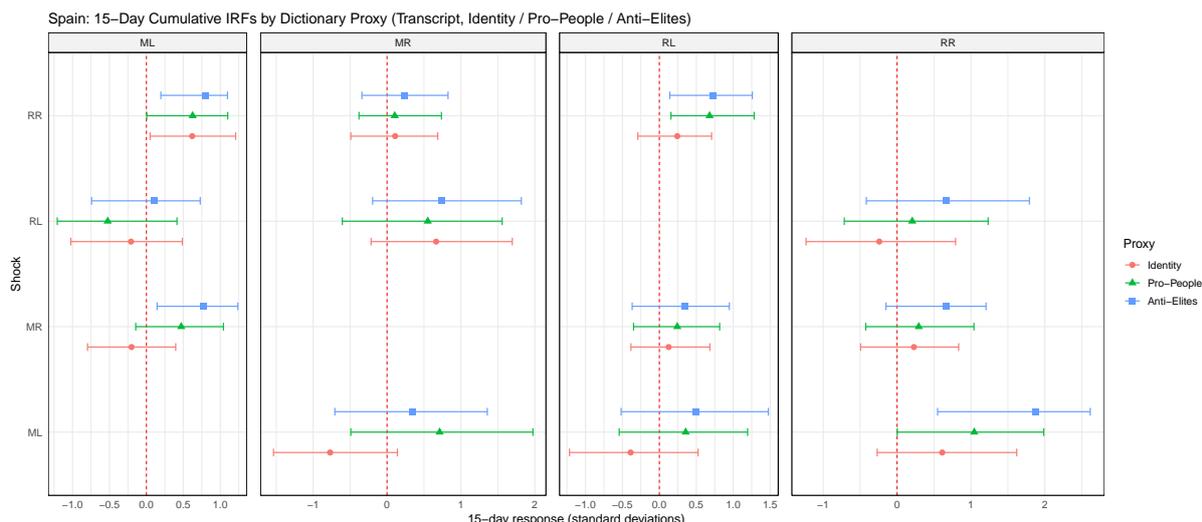
(a) Spain

(b) United Kingdom

Figure 11: VAR Results for Dictionary-Based Populist Discourse Measures in Spain and the United Kingdom

Notes: Panel (a) reports Spain and panel (b) the United Kingdom. The dependent variables are transcript-based dictionary measures of identity, pro-people rhetoric, and anti-elite rhetoric. The figures display cross-measure 15-day cumulative impulse-response estimates from separate VAR(4) models, excluding own-response terms. The underlying measures are dictionary-based indicators of identity, pro-people rhetoric, and anti-elite rhetoric constructed from YouTube transcripts. In each specification, each measure enters as both a dependent variable and a lagged predictor, and election-period dummies are included. Points denote standardized effects and horizontal bars indicate bootstrap confidence intervals. The estimation sample includes $N_{\text{Spain}} = 1,483$ daily observations and $N_{\text{UK}} = 2,182$ daily observations.

United Kingdom between 2015 and 2024, we developed temporally disaggregated, multi-dimensional measures of party rhetoric that make it possible to trace campaign interaction as it unfolds.

Methodologically, we introduce a general framework for measuring campaign communication with high-frequency social media data. By integrating embeddings, large language models, BERTopic, and dictionaries within a unified design, the framework moves beyond single representations of text and captures multiple dimensions of political communication—issue attention, ideological positioning, and

rhetorical proximity—within the same empirical setting. This design makes it possible to distinguish whether observed changes reflect shifts in what parties talk about, how they position themselves, or how they frame similar issues, and to analyze these dimensions jointly over time.

Our findings show that parties do not simply converge under electoral pressure as election day approaches. Instead, rhetorical adaptation is strategic, asymmetric, and conditioned by the structure of competition. In the United Kingdom, where the central contest for government remains one between two mainstream parties, campaign dynamics are largely mainstream-led. Outsider parties exert influence mainly through issue differentiation and selective rightward movement rather than by setting the broader terms of campaign debate. Issue attention increasingly reflects core electoral identities, limiting cross-party agenda-setting. In Spain, by contrast, responsiveness is structured above all within the left bloc, where coalition prospects have been most credible since 2015. As elections approach, both the mainstream left and the radical left move further left ideologically, party rhetoric departs from manifesto anchors across most party families, and campaign discourse shifts away from substantive policy debate toward partisan contestation and governing arrangements. Shocks affecting the mainstream left are followed by movements in the radical left, and vice versa, in a pattern consistent with coalition-building dynamics.

These findings speak to broader debates on mainstream and outsider parties. Much existing work has treated campaign influence as running mainly from outsiders to mainstream parties, especially on highly politicized issues such as immigration. Our results suggest a more conditional view. Campaign influence is neither uniformly outsider-led nor fully reciprocal. In some contexts, outsiders may be the ones following mainstreams.

These findings may travel beyond the two cases examined here. The Spanish pattern may extend to other multiparty systems in which coalition competition structures interaction between mainstream and outsider parties, whether outsiders are present on both sides of the ideological spectrum, as in Greece, or primarily on one side, as in Germany. The British pattern may be more relevant to majoritarian systems in which outsider parties shape competition without displacing mainstream parties as the principal contenders for office. The analysis is also limited to YouTube. Yet, YouTube has become an increasingly important channel for news and political communication (Schomer 2020; Konitzer et al. 2020; Hosseinmardi et al. 2021), and campaign videos are frequently recirculated across platforms such as X, Facebook, and Instagram; hence, the patterns documented here are unlikely to be specific to a single platform.

More broadly, this study shows how social media data can be used to recover parties' short-term strategic behavior during campaigns. It also shows that recent NLP tools can capture not only what parties discuss, but also where they place themselves ideologically and how closely their rhetoric aligns with

that of rivals and allies at campaign speed. Taken together, these contributions make it possible to study party competition at a level of temporal and substantive detail that traditional campaign data rarely permit. Extending this approach to local elections would be especially useful, since local campaigns operate under different institutional incentives, levels of media attention, and organizational constraints. Comparing local and national contests would help clarify when the patterns identified here travel across levels of representation and when they do not.

REFERENCES

Abou-Chadi, Tarik. 2016. "Niche Party Success and Mainstream Party Policy Shifts – How Green and Radical Right Parties Differ in Their Impact." *British Journal of Political Science* 46 (2): 417–436.

Abou-Chadi, Tarik, Christoffer Green-Pedersen, and Peter B. Mortensen. 2020. "Parties' policy adjustments in response to changes in issue saliency." *West European Politics* 43 (4): 749–771.

Abou-Chadi, Tarik, and Werner Krause. 2020. "The Causal Effect of Radical Right Success on Mainstream Parties' Policy Positions: A Regression Discontinuity Approach." *British Journal of Political Science* 50 (3): 829–847.

Abou-Chadi, Tarik, and Markus Wagner. 2019. "The electoral appeal of party strategies in postindustrial societies: when can the mainstream left succeed?" *The Journal of Politics* 81 (4): 1405–1419.

Adams, James, and Zeynep Somer-Topcu. 2009. "Policy Adjustment by Parties in Response to Rival Parties' Policy Shifts: Spatial Theory and the Dynamics of Party Competition in Twenty-Five Post-War Democracies." *British Journal of Political Science* 39 (4): 825–846.

Akkerman, Tjitske. 2015. "Immigration policy and electoral competition in Western Europe: A fine-grained analysis of party positions over the past two decades." *Party Politics* 21 (1): 54–67.

Alonso, Sonia, and Sara Claro da Fonseca. 2012. "Immigration, left and right." *Party Politics* 18 (6): 865–884.

Bale, Tim, Christoffer Green-Pedersen, AndréA Krouwel, Kurt Richard Luther, and Nick Sitter. 2010. "If you can't Beat them, Join them? Explaining Social Democratic Responses to the Challenge from the Populist Radical Right in Western Europe." *Political Studies* 58 (3): 410–426.

Barberá, Pablo, Andreu Casas, Jonathan Nagler, Patrick J. Egan, Richard Bonneau, John T. Jost, and Joshua A. Tucker. 2019. "Who Leads? Who Follows? Measuring Issue Attention and Agenda Setting by Legislators and the Mass Public Using Social Media Data." *American Political Science Review* 113 (4): 883–901.

Berz, Jan, and Corinna Kroeber. 2023. "Walking the line: Electoral cycles and the shift in legislative priorities among German parliamentarians." *Electoral Studies* 83:102595.

Bestvater, Samuel E., and Burt L. Monroe. 2023. "Sentiment is not stance: Target-aware opinion classification for political text analysis." *Political Analysis* 31 (2): 235–256.

Bond, Robert, and Solomon Messing. 2015. "Quantifying social media's political space: Estimating ideology from publicly revealed preferences on Facebook." *American Political Science Review* 109 (1): 62–78.

Brandt, Patrick T., Michael Colaresi, and John R. Freeman. 2008. "The Dynamics of Reciprocity, Accountability, and Credibility." *Journal of Conflict Resolution* 52 (3): 343–374.

Brandt, Patrick T., and John R. Freeman. 2009. "Modeling Macro-Political Dynamics." *Political Analysis* 17 (2): 113–142.

Bräuninger, Thomas, and Nathalie Giger. 2018. "Strategic ambiguity of party positions in multi-party competition." *Political Science Research and Methods* 6 (3): 527–548.

Conover, Michael, Jacob Ratkiewicz, Matthew Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. 2011. "Political polarization on twitter." In *Proceedings of the international aaai conference on web and social media,* 5:89–96.

Cowburn, Mike, and Marius Sältzer. 2025. "Partisan communication in two-stage elections: the effect of primaries on intra-campaign positional shifts in congressional elections." *Political Science Research and Methods* 13 (2): 392–411.

Dalton, Russell J., and Ian McAllister. 2015. "Random Walk or Planned Excursion? Continuity and Change in the Left–Right Positions of Political Parties." *Comparative Political Studies* 48 (6): 759–787.

Del Vicario, Michela, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. 2016. "Echo chambers: Emotional contagion and group polarization on facebook." *Scientific reports* 6 (1): 37825.

Dennison, James. 2016. *The Greens in British politics: protest, anti-austerity and the divided left.* Springer.

Edwards, George C., and B. Dan Wood. 1999. "Who influences whom? The president, Congress, and the media." *American Political Science Review* 93 (2): 327–344.

Enders, Walter, and Todd Sandler. 1993. "The effectiveness of antiterrorism policies: A vector-autoregression-intervention analysis." *American Political Science Review* 87 (4): 829–844.

Freeman, John R., John T. Williams, and Tse-min Lin. 1989. "Vector autoregression and the study of politics." *American Journal of Political Science,* 842–877.

Gessler, Theresa, and Sophia Hunger. 2022. "How the refugee crisis and radical right parties shape party competition on immigration." *Political Science Research and Methods* 10 (3): 524–544.

Gilardi, Fabrizio, Theresa Gessler, Maël Kubli, and Stefan Müller. 2021. "Social Media and Policy Responses to the COVID-19 Pandemic in Switzerland." *Swiss Political Science Review* 27 (2): 243–256.

————. 2022. "Social Media and Political Agenda Setting." *Political Communication* 39 (1): 39–60.

Gonzalez-Rostani, Valentina. 2026. "Elections, Right-Wing Populism, and Political-Economic Polarization: The Role of Institutions and Political Outsiders." *The Journal of Politics* 88 (1): 125–144.

González-Rostani, Valentina, José Incio, and Guillermo Lezama. 2025. "Social media versus surveys: A new scalable approach to understanding legislators' discourse." *Legislative Studies Quarterly* 50 (2): 258–266.

Green-Pedersen, Christoffer, and Simon Otjes. 2019. "A hot topic? Immigration on the agenda in Western Europe." *Party Politics* 25 (3): 424–434.

Grootendorst, Maarten. 2022. *BERTopic: Neural topic modeling with a class-based TF-IDF procedure.*

Gruber, Oliver. 2014. *Campaigning in Radical Right Heartland: The politicization of immigration and ethnic relations in Austrian general elections, 1971-2013.* LIT Verlag Münster.

Han, Kyung Joon. 2015. "The Impact of Radical Right-Wing Parties on the Positions of Mainstream Parties Regarding Multiculturalism." *West European Politics* 38 (3): 557–576.

Haroon, Muhammad, Magdalena Wojcieszak, Anshuman Chhabra, Xin Liu, Prasant Mohapatra, and Zubair Shafiq. 2023. "Auditing YouTube's recommendation system for ideologically congenial, extreme, and problematic recommendations." *Proceedings of the National Academy of Sciences* 120 (50): e2213020120.

Heerden, Sjoerdje van, Sarah L. de Lange, Wouter van der Brug, and Meindert Fennema. 2014. "The Immigration and Integration Debate in the Netherlands: Discursive and Programmatic Reactions to the Rise of Anti-Immigration Parties." *Journal of Ethnic and Migration Studies* 40 (1): 119–136.

Hopmann, David N., Christian Elmelund-Præstekær, Erik Albæk, Rens Vliegenthart, and Claes H. de Vreese. 2012. "Party media agenda-setting: How parties influence election news coverage." *Party Politics* 18 (2): 173–191.

Hosseinmardi, Homa, Amir Ghasemian, Aaron Clauset, Markus Mobius, David M. Rothschild, and Duncan J. Watts. 2021. "Examining the consumption of radical content on YouTube." *Proceedings of the National Academy of Sciences* 118 (32): e2101967118.

Hosseinmardi, Homa, Amir Ghasemian, Miguel Rivera-Lanas, Manoel Horta Ribeiro, Robert West, and Duncan J. Watts. 2024. "Causally estimating the effect of YouTube's recommender system using counterfactual bots." *Proceedings of the National Academy of Sciences* 121 (8): e2313377121.

Hutter, Swen, and Hanspeter Kriesi. 2022. "Politicising immigration in times of crisis." *Journal of Ethnic and Migration Studies* 48 (2): 341–365.

Ibrahim, Hazem, Nouar AlDahoul, Sangjin Lee, Talal Rahwan, and Yasir Zaki. 2023. "YouTube's recommendation algorithm is left-leaning in the United States." *PNAS Nexus* 2 (8): pgad264.

Iversen, Torben, and David Soskice. 2006. "Electoral institutions and the politics of coalitions: Why some democracies redistribute more than others." *American political science review* 100 (2): 165–181.

Klüver, Heike, and Iñaki Sagarzazu. 2016. "Setting the Agenda or Responding to Voters? Political Parties, Voters and Issue Attention." *West European Politics* 39 (2): 380–398.

Konitzer, T., J. Allen, S. Eckman, B. Howland, M. M. Mobius, D. M. Rothschild, and D. Watts. 2020. "Measuring news consumption with behavioral versus survey data." *SSRN,* 10.

Lai, Angela, Megan A. Brown, James Bisbee, Joshua A. Tucker, Jonathan Nagler, and Richard Bonneau. 2024. "Estimating the Ideology of Political YouTube Videos." *Political Analysis* 32 (3): 345–360.

Laurer, Moritz, Wouter van Atteveldt, Andreu Casas, and Kasper Welbers. 2024. "Less Annotating, More Classifying: Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI." *Political Analysis* 32 (1): 84–100.

Le Pennec, Caroline. 2024. "Strategic campaign communication: Evidence from 30,000 candidate manifestos." *The Economic Journal* 134 (658): 785–810.

Licht, Hauke. 2023. "Cross-lingual classification of political texts using multilingual sentence embeddings." *Political Analysis* 31 (3): 366–379.

Licht, Hauke, Tarik Abou-Chadi, Pablo Barberá, and Whitney Hua. 2025. "Measuring and Understanding Parties' Anti-elite Strategies." *The Journal of Politics* 87:275–290.

Linegar, Mitchell, Rafal Kocielnik, and R. Michael Alvarez. 2023. "Large language models and political science." *Frontiers in Political Science* 5.

Mamié, Robin, Manoel Horta Ribeiro, and Robert West. 2021. "Are Anti-Feminist Communities Gateways to the Far Right? Evidence from Reddit and YouTube." In *13th ACM Web Science Conference 2021,* 139–147. Virtual Event United Kingdom: ACM.

Meguid, Bonnie M. 2005. "Competition between Unequals: The Role of Mainstream Party Strategy in Niche Party Success." *The American Political Science Review* 99 (3): 347–359.

Mens, Gaël Le, and Aina Gallego. 2025. "Positioning Political Texts with Large Language Models by Asking and Averaging." *Political Analysis* 33 (3): 274–282.

Meyer, Sarah, and Sieglinde Rosenberger. 2015. "Just a shadow? The role of radical right parties in the politicization of immigration, 1995–2009." *Politics and Governance* 3 (2): 1–17.

Meyer, Thomas M., and Markus Wagner. 2021. "Issue Engagement Across Members of Parliament: The Role of Issue Specialization and Party Leadership." *Legislative Studies Quarterly* 46 (3): 653–678.

Mohsin, Maryam. 2020. *YouTube stats every marketer should know in 2020.* Technical report.

Poljak, Željko, and Annelise Russell. 2025. "All Talk, No Action? Politicians' Agenda Responsiveness to Citizens' Engagement on Social Media." *Political Research Quarterly* 78 (2): 495–508.

Polk, Jonathan, Jan Rovny, Ryan Bakker, Erica Edwards, Liesbet Hooghe, Seth Jolly, Jelle Koedam, et al. 2017. "Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data." *Research & Politics* 4 (1): 205316801668691.

Qin, Duo. 2011. "Rise of VAR modelling approach." *Journal of Economic Surveys* 25 (1): 156–174.

Rheault, Ludovic, and Christopher Cochrane. 2020. "Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora." *Political Analysis* 28 (1): 112–133.

Rodriguez, Pedro L., Arthur Spirling, and Brandon M. Stewart. 2023. "Embedding Regression: Models for Context-Specific Description and Inference." *American Political Science Review* 117 (4): 1255–1274.

Rooduijn, Matthijs, and Teun Pauwels. 2011. "Measuring Populism: Comparing Two Methods of Content Analysis." *West European Politics* 34 (6): 1272–1283.

Ruedin, Didier, and Laura Morales. 2019. "Estimating party positions on immigration: Assessing the reliability and validity of different methods." *Party Politics* 25 (3): 303–314.

Sagarzazu, Iñaki, and Heike Klüver. 2017. "Coalition Governments and Party Competition: Political Communication Strategies of Coalition Parties*." *Political Science Research and Methods* 5 (2): 333–349.

Scherr, Sebastian, Carsten Reinemann, and Olaf Jandura. 2015. "Dynamic Success on YouTube: A Longitudinal Analysis of Click Counts and Contents of Political Candidate Clips during the 2009 German National Election." *German Politics* 24 (4): 491–506.

Schomer, Audrey. 2020. *US YouTube Advertising 2020.* Technical report.

Seeberg, Henrik Bech. 2022. "First avoidance, then engagement: Political parties' issue competition in the electoral cycle." *Party Politics* 28 (2): 284–293.

Sohal, Seerat, and Harsandaldeep Kaur. 2018. "A Content Analysis of YouTube Political Advertisements: Evidence from Indian Parliamentary Elections." *Journal of Creative Communications* 13 (2): 133–156.

Somer-Topcu, Zeynep. 2009. "Timely Decisions: The Effects of Past National Elections on Party Policy Change." *The Journal of Politics* 71 (1): 238–248.

Van Spanje, Joost. 2010. "Contagious Parties: Anti-Immigration Parties and Their Impact on Other Parties' Immigration Stances in Contemporary Western Europe." *Party Politics* 16 (5): 563–586.

Vesnic-Alujevic, Lucia, and Sofie Van Bauwel. 2014. "YouTube: A Political Advertising Tool? A Case Study of the Use of YouTube in the Campaign for the European Parliament Elections 2009." *Journal of Political Marketing* 13 (3): 195–212.

Wagner, Markus, and Thomas M. Meyer. 2014. "Which Issues do Parties Emphasise? Salience Strategies and Party Organisation in Multiparty Systems." *West European Politics* 37 (5): 1019–1045.

Widmann, Tobias, and Maximilian Wich. 2023. "Creating and comparing dictionary, word embedding, and transformer-based models to measure discrete emotions in German political text." *Political Analysis* 31 (4): 626–641.

Williams, Christopher J., and Sophia Hunger. 2022. "How challenger party issue entrepreneurship and mainstream party strategies drive public issue salience: evidence from radical-right parties and the issue of immigration." *European Political Science Review* 14 (4): 544–565.

Wirsching, Elisa M., Pedro L. Rodriguez, Arthur Spirling, and Brandon M. Stewart. 2025. "Multilanguage Word Embeddings for Social Scientists: Estimation, Inference, and Validation Resources for 157 Languages." *Political Analysis* 33 (2): 156–163.

Online Appendix

Contents

A Youtube Data

A.1 Descriptives Youtube Data

TABLE A.1.1: Descriptives Youtube Data by Party Family

	Spain			United Kingdom	
Family	Party	Total videos	Family	Party	Total videos
ML	PSOE	1693	ML	Green	374
MR	PP	502	ML	Labour	100
MR	Ciudadanos	988	MR	Conservative	307
RL	Izquierda Unida	1333	MR	Liberal Democrats	559
RL	Podemos	859	RR	Reform UK	123
RL	Sumar	95	RR	UKIP	400
RR	VOX	1420			
Total		6890	Total		1863

TABLE A.1.2: Descriptives Youtube Data by Party Family and Election Year

	Spain			United Kingdom	
Family	Election Year	Total videos	Family	Election Year	Total videos
ML	2015	420	ML	2015	142
ML	2016	417	ML	2016 (Brexit referendum)	62
ML	2019 (April)	468	ML	2017	70
ML	2019 (November)	14	ML	2019	30
ML	2023	374	ML	2021	119
MR	2015	258	ML	2024	51
MR	2016	260	MR	2015	225
MR	2019 (April)	155	MR	2016 (Brexit referendum)	74
MR	2019 (November)	179	MR	2017	260
MR	2023	638	MR	2019	78
RL	2015	546	MR	2021	81
RL	2016	452	MR	2024	148
RL	2019 (April)	619	RR	2015	48
RL	2019 (November)	244	RR	2016 (Brexit referendum)	36
RL	2023	426	RR	2017	51
RR	2015	151	RR	2019	193
RR	2016	108	RR	2021	30
RR	2019 (April)	548	RR	2024	165
RR	2019 (November)	51			
RR	2023	562			
Total		6890	Total		1863

TABLE A.1.3: Descriptives Youtube Data by Election Year, Country and Party

Country	Election Year	Party	Total videos	With transcript	Missing or ≤5 words	Usable transcript	Share with transcript (%)	Share missing or ≤5 words (%)
Spain	2015	Ciudadanos	258	180	78	180	69.77	30.23
Spain	2015	Izquierda Unida	421	385	36	385	91.45	8.55
Spain	2015	PSOE	420	374	47	373	89.045	11.19
Spain	2015	Podemos	125	108	18	107	86.40	14.40
Spain	2015	VOX	151	128	28	123	84.77	18.54
Spain	2016	Ciudadanos	260	194	66	194	74.62	25.38
Spain	2016	Izquierda Unida	179	162	17	162	90.50	9.50
Spain	2016	PSOE	417	369	49	368	88.49	11.75
Spain	2016	Podemos	273	212	64	209	77.66	23.44
Spain	2016	VOX	108	98	10	98	90.74	9.26
Spain	2019A	Ciudadanos	123	81	42	81	65.85	34.15
Spain	2019A	Izquierda Unida	392	229	163	229	58.42	41.58
Spain	2019A	PP	32	23	11	21	71.88	34.38
Spain	2019A	PSOE	468	312	177	291	0.67	0.38
Spain	2019A	Podemos	227	156	72	155	68.72	31.72
Spain	2019A	VOX	548	429	120	428	78.28	21.90
Spain	2019B	Ciudadanos	175	103	72	103	58.86	41.14
Spain	2019B	Izquierda Unida	135	108	27	108	80	20
Spain	2019B	PP	4	1	3	1	25	75
Spain	2019B	PSOE	14	7	7	7	50	50
Spain	2019B	Podemos	109	86	23	86	78.90	21.10
Spain	2019B	VOX	51	37	14	37	72.55	27.45
Spain	2023	Ciudadanos	172	162	10	162	94.19	5.81
Spain	2023	Izquierda Unida	206	204	2	204	99.03	0.97
Spain	2023	PP	466	441	26	440	94.63	5.58
Spain	2023	PSOE	374	341	35	339	91.18	9.36
Spain	2023	Podemos	125	116	9	116	92.80	7.20
Spain	2023	Sumar	95	74	21	74	77.89	22.11
Spain	2023	VOX	562	535	29	533	95.16	5.16
UK	2015	Conservative	105	96	10	95	91.43	9.52
UK	2015	Green	142	107	35	107	75.35	24.65
UK	2015	Liberal Democrats	120	116	5	115	96.67	4.17
UK	2015	UKIP	48	48	0	48	100	0
UK	2016	Conservative	51	50	1	50	98.04	1.96
UK	2016	Green	62	62	0	62	100	0
UK	2016	Liberal Democrats	23	23	0	23	100	0
UK	2016	UKIP	36	36	0	36	100	0
UK	2017	Conservative	53	51	2	51	96.23	3.77
UK	2017	Green	61	58	3	58	95.08	4.92
UK	2017	Labour	9	9	0	9	100	0
UK	2017	Liberal Democrats	207	187	20	187	90.34	9.66
UK	2017	UKIP	51	50	1	50	98.04	1.96
UK	2019	Conservative	41	38	3	38	92.68	7.32
UK	2019	Green	21	18	3	18	85.71	14.29
UK	2019	Labour	9	9	0	9	100	0
UK	2019	Liberal Democrats	37	30	7	30	81.08	18.92
UK	2019	UKIP	193	187	6	187	96.89	3.11
UK	2021	Conservative	14	14	0	14	100	0
UK	2021	Green	62	62	0	62	100	0
UK	2021	Labour	57	53	4	53	92.98	0.070175
UK	2021	Liberal Democrats	67	64	4	63	95.52	5.97
UK	2021	UKIP	30	29	1	29	96.67	3.33
UK	2024	Conservative	43	42	1	42	97.67	02.33
UK	2024	Green	26	26	0	26	100	0
UK	2024	Labour	25	21	5	20	84	20
UK	2024	Liberal Democrats	105	94	11	94	89.52	10.48
UK	2024	ReformUK	123	121	3	120	98.37	2.44
UK	2024	UKIP	42	42	0	42	100	0

A.2 YouTube Data Collection Steps

Data collection was conducted using Python in two main stages: first, identifying and retrieving video metadata from official party channels; second, obtaining and processing the available transcripts. The following steps outline the methodology used:

1. **Building the YouTube API Client.** We built the YouTube API client using the `googleapiclient.discovery` module, which allows for efficient retrieval of video metadata from specified channels via the YouTube Data API.

2. **Retrieving Video Details.** For each channel, we retrieved the title, publication date, and URL of all videos posted within the defined date ranges corresponding to the year before each electoral event.

3. **Fetching Video Transcripts.** We obtained the available auto-generated transcripts for each identified video using the `YouTubeTranscriptApi`, which extracts textual content directly from YouTube's captioning system.

4. **Language Filtering.** Transcripts were filtered based on the language of the source video. The `YouTubeTranscriptApi` supports language-specific retrieval via language codes, and we used this functionality to retrieve transcripts in their original language– English for UK videos and Spanish for Spanish videos.

5. **Translating Spanish Transcripts.** Transcripts retrieved in Spanish were translated into English using the Google Translate API, accessed via the `googletrans` Python library. This step ensured that all transcripts were available in a common language for downstream processing and cross-country comparability.

Following translation and data cleaning, to ensure analytic quality, we retained only transcripts containing more than five words after cleaning. This threshold was applied to exclude transcripts that consisted of non-coherent or minimal text, a common occurrence when YouTube's auto-captioning system captures background noise, music, or other non-speech audio. Out of the 8,753 videos collected, 7,398 yielded usable transcripts meeting this criterion and constitute the final analytic sample.

B Measuring Ideology

B.1 Prompt and Model Specification

We estimated the ideological position of each transcript using OpenAI's `gpt-4o-mini` model via the Responses API. In the API call, the instruction text below was passed in the `instructions` field, while the transcript itself was passed separately as the `input`. The prompt was used exactly as follows:

> You are a classifier that scores the ideological position of political campaign content. You will be given the text of a YouTube campaign-related video transcript from parties in Spain or the UK between 2015 and 2023. Text will appear in Spanish or English Where does this text stand on the 'left' to 'right' wing scale? Determine where the text stands on a left–right political scale. Return an integer score from 0 to 100 (inclusive), where:
> - 0 = Extremely left.
> - 50 = Center.
> - 100 = Extremely right.
> If the text contains no meaningful political or ideological content, return 'NA'. Only return one of the scores between 0 to 100. If the content is mixed or ambiguous but political, choose the single best overall score. Do not provide explanations, and do not repeat or transform the input text. Do not explain your reasoning or return any other text.

The model was run with `temperature = 0`. This setting minimizes randomness in generation and was chosen to maximize consistency and reproducibility across transcripts. The model was instructed to return a single integer from 0 to 100, where 0 denotes the extreme left, 50 the center, and 100 the extreme right. If the transcript contained no meaningful political or ideological content, the model was instructed to return `NA`.

After generation, outputs were post-processed so that only integers in the range `[0,100]` or `NA` were retained. Any other output was treated as an error. Before scoring, transcript text was stripped of leading and trailing whitespace, repeated whitespace was collapsed, and texts were truncated to a preset maximum character length where necessary.
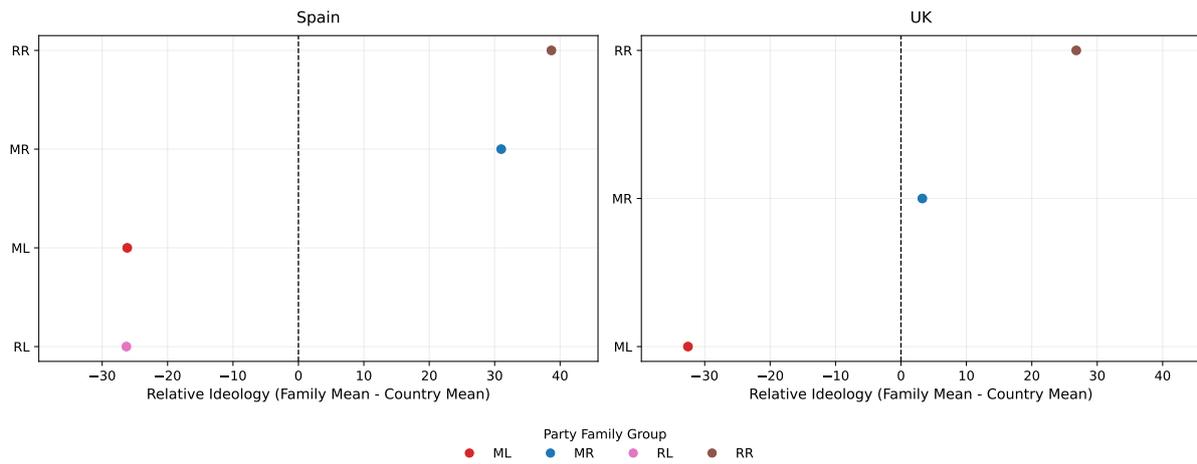
B.2 Descriptives Ideology

Figure B.2.1: Relative Average Levels of Ideology by Party Family and Country

Notes: Points indicate each party family's average transcript-based AI ideology score relative to the country average within Spain or the UK. Relative ideology is calculated as the family mean minus the country mean: positive values indicate positions above the country average, while negative values indicate positions below the country average. The underlying ideology scores range from 0 (extreme left) to 100 (extreme right). Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right.

C Measuring Issue Attention: BERTopic
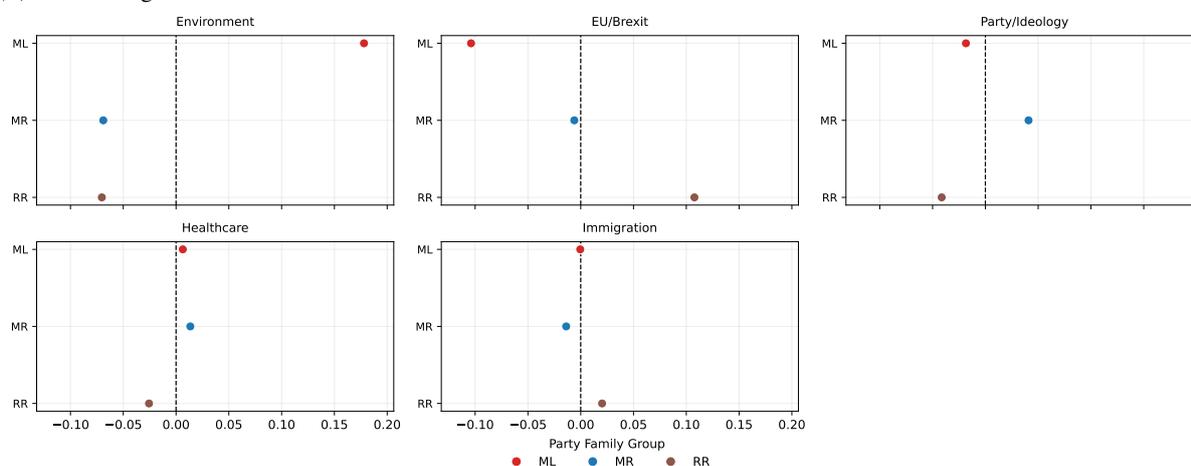
C.1 Descriptives BERTopic

(a) Spain
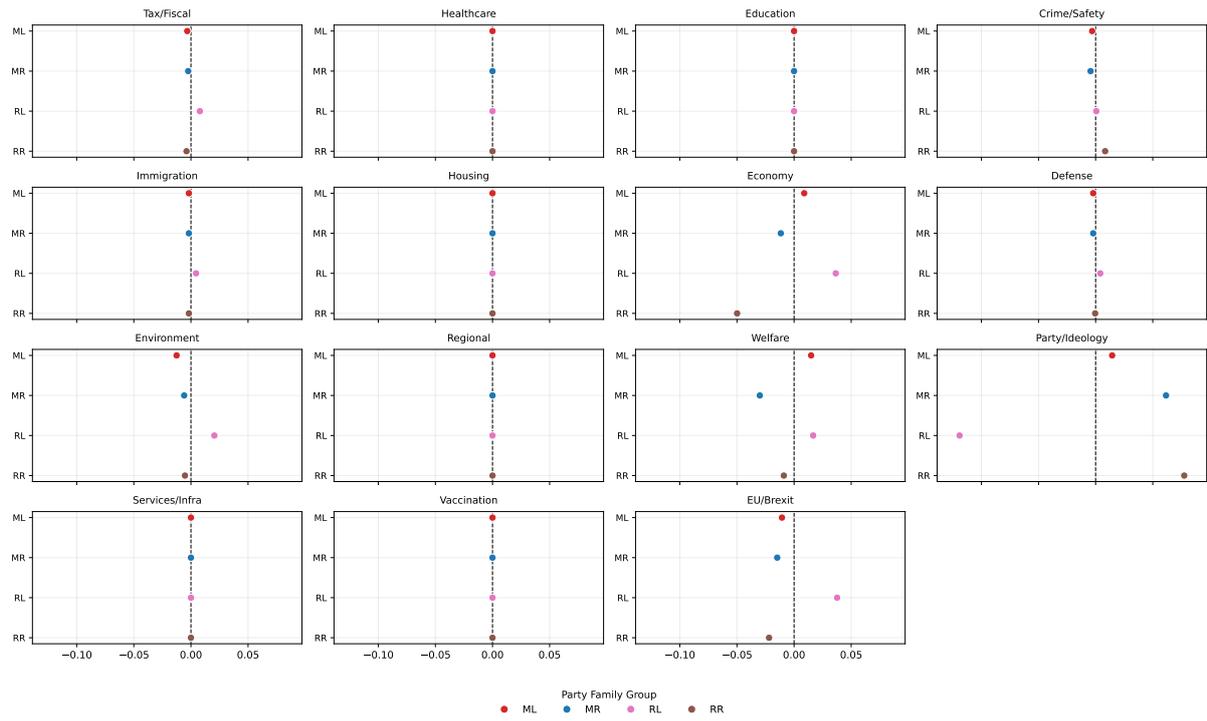
(b) United Kingdom
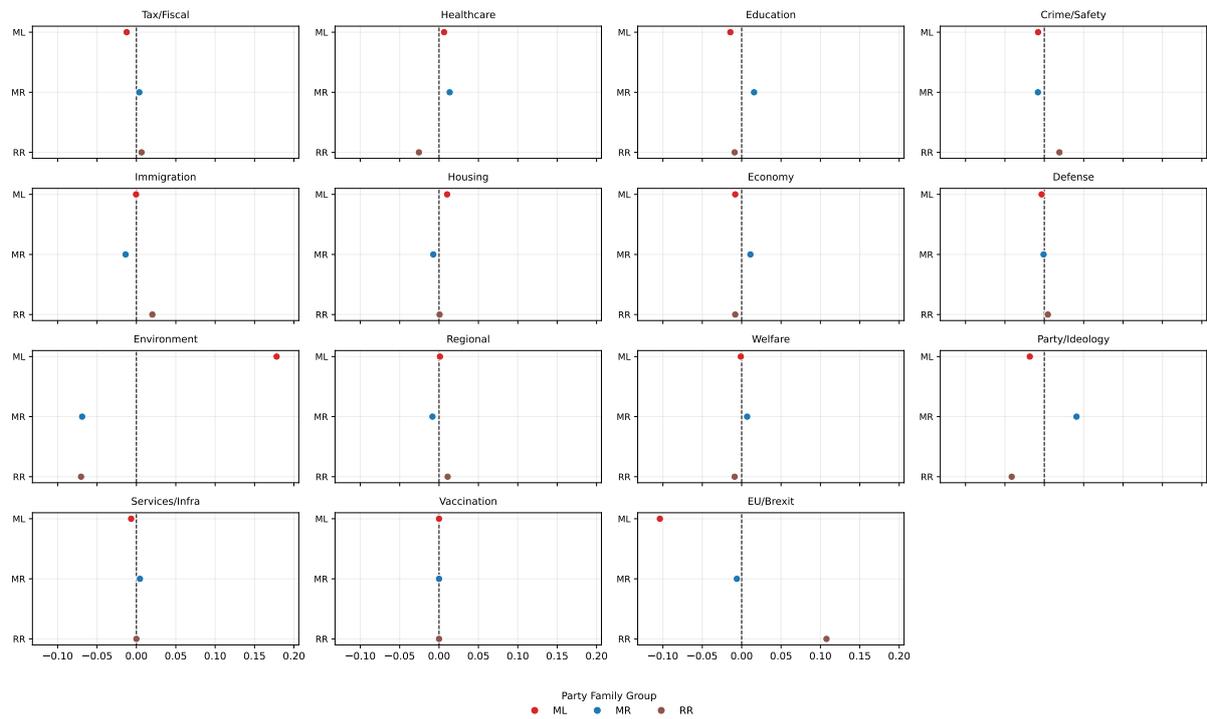
Figure C.1.1: Relative Average BERTopic Topic Shares by Party Family in Spain and the United Kingdom

Notes: Panel (a) reports Spain and panel (b) the United Kingdom. Each small panel corresponds to a BERTopic dimension derived from YouTube transcripts. Points indicate each party family's average transcript-based topic share relative to the overall country average for that topic. Relative topic share is calculated as the party-family mean minus the country mean, so positive values indicate greater emphasis on the topic than the country average and negative values indicate less emphasis. Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right. The UK panels exclude the radical-left category because no UK parties in the sample are classified in that family.

(a) Spain
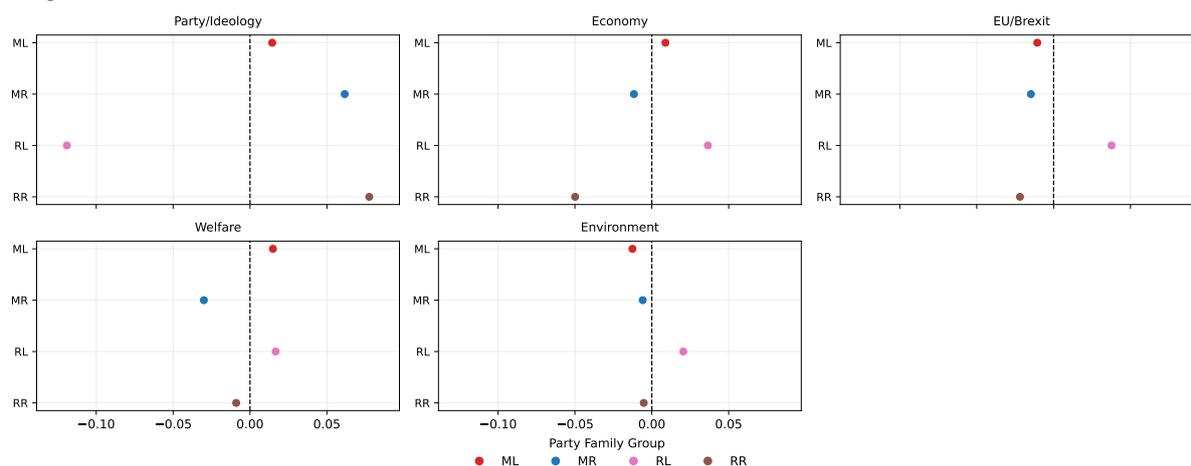
(b) United Kingdom

Figure C.1.2: Relative Average BERTopic Topic Shares for the Top 5 Topics by Party Family in Spain and the United Kingdom

Notes: Panel (a) reports Spain and panel (b) the United Kingdom. Each small panel corresponds to one of the five BERTopic dimensions with the largest absolute family-level deviation from the country average topic share in that country, based on YouTube transcripts. Points indicate each party family's average transcript-based topic share relative to the overall country average for that topic. Relative topic share is calculated as the party-family mean minus the country mean, so positive values indicate greater emphasis on the topic than the country average and negative values indicate less emphasis. Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right. The UK panels exclude the radical-left category because no UK parties in the sample are classified in that family.

D MEASURING ISSUE ATTENTION: DICTIONARIES

D.1 Identity Dictionary

The identity dictionary captures references to cultural identity, national way of life, and nativist rhetoric, organized into four subcategories:

- *Immigration and border politics:* immigr, border, wall, undocumented, illega, asyl, migrat, migrant, migrier, deport, enforcement.

- *Cultural identity and national way of life:* heritage, values, culture, way of life, tradition, inclusion.

- *Religion and sexuality-based identity politics:* muslim, christian, islam, gay, lesbian, lgbt.

- *Security and threat frames:* terrorism, xenophob, invasion.

D.2 Anti-Elite Dictionary

The anti-elite dictionary captures rhetoric directed against political and social elites, organized into three subcategories:

- *Class and establishment frames:* cast, class, elit, elitist, establish, polit, politic, politician.

- *Corruption and authoritarian frames:* corrupt, regim, regimen, rule, propaganda, directori.

- *Moral condemnation and dismissal frames:* promin, arrog, arrogantli, betrai, treason, promis, shame, undemocrat, deceit, absurd, absurdli, admit, admitt.

D.3 Pro-People Dictionary

The pro-people dictionary captures rhetoric appealing directly to ordinary people and participatory democracy:

- peopl, tradit, tradition, direct, directli, referendum.
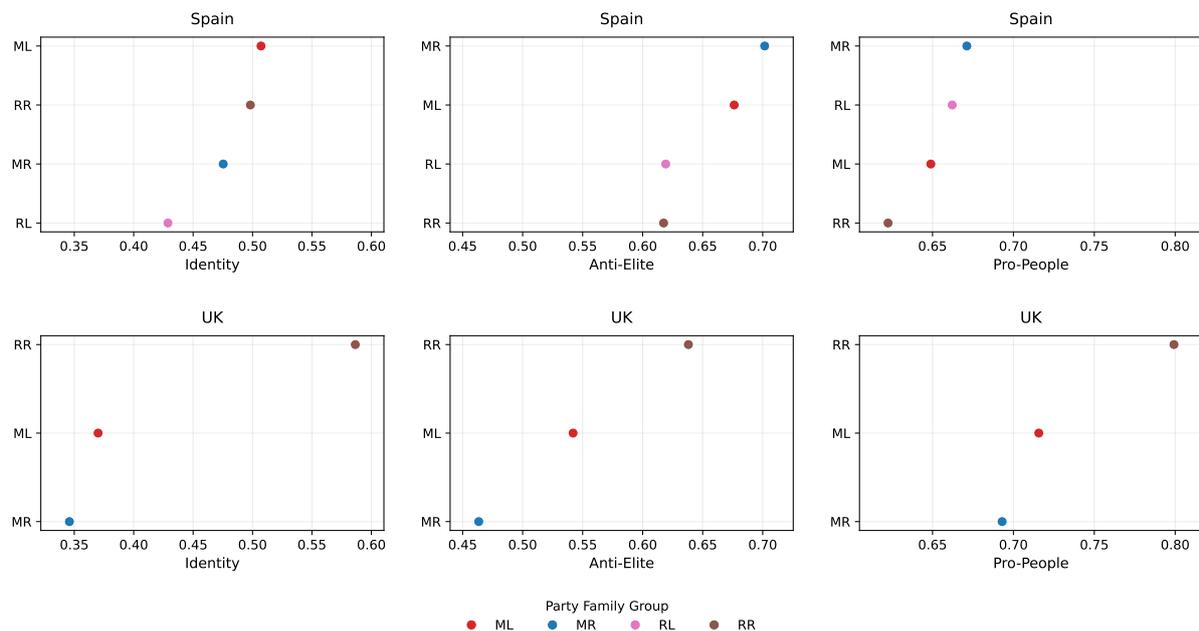
D.4 Descriptives Dictionaries

Figure D.4.1: Average Levels of Identity, Anti-Elite Rhetoric, and Pro-People Rhetoric by Party Family and Country

Notes: Points indicate the average transcript-based dictionary count for each party family within each country. The three panels show average levels of identity, anti-elite rhetoric, and pro-people rhetoric. Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right.
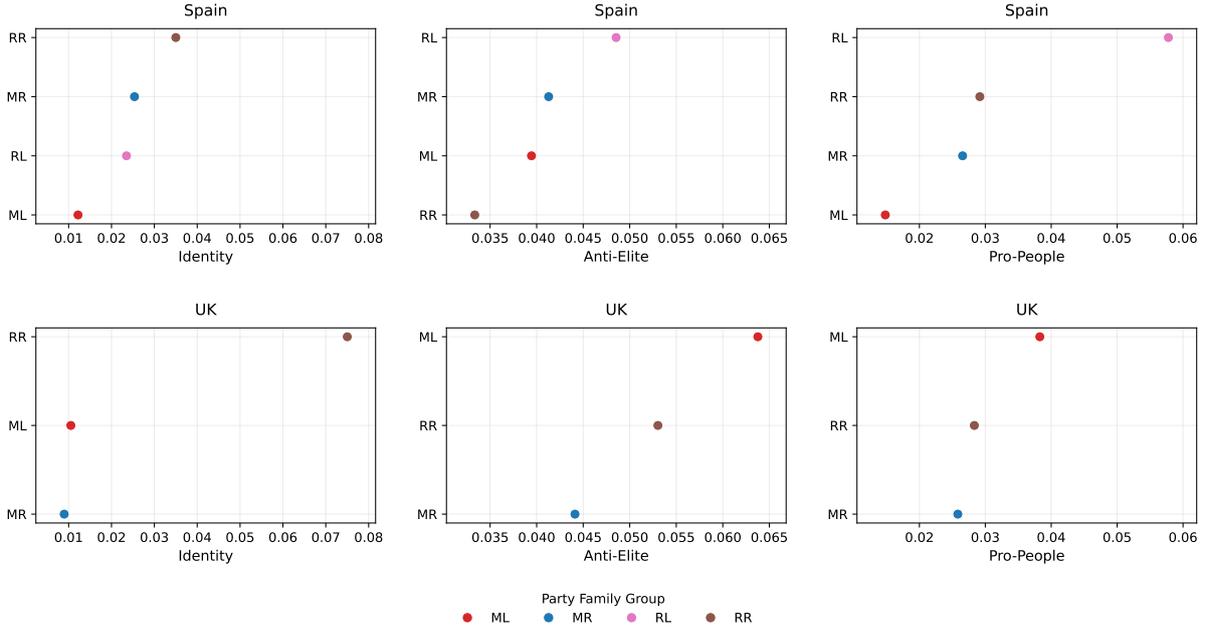
Figure D.4.2: Relative Average Levels of Identity, Anti-Elite Rhetoric, and Pro-People Rhetoric by Party Family and Country

Notes: Points indicate the average relative transcript-based dictionary score for each party family within each country. The three panels show average levels of identity, anti-elite rhetoric, and pro-people rhetoric. Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right.

E ANALYSIS: CLOSE TO ELECTIONS ANALYSIS

E.1 Alternative Specification: Close-to-Election Binary Indicator

As a robustness check, we re-estimate the election-proximity models replacing the continuous days-to-election measure with a binary indicator. Specifically, $close_i = 1$ if the video was posted within 30 calendar days of election day and 0 otherwise. The estimating equation is:

$$y_{im} = \alpha_m + \beta_m \cdot close_i + \sum_{f \neq f_0} \gamma_{fm} D_{if} + \sum_{f \neq f_0} \delta_{fm}(close_i \cdot D_{if}) + \sum_{e \neq e_0} \eta_{em} P_{ie} + \theta_m \cdot UK_i + \varepsilon_{im}$$

where y_{im} denotes the outcome of interest m for observation i, including LLM-based ideology scores, dictionary measures of identity and populist discourse, and BERTopic issue shares. The coefficient β_m captures the average difference in outcome m during the final 30-day window relative to the rest of the campaign year for the reference party family, and the interaction terms δ_{fm} recover family-specific deviations from this baseline. Election-period fixed effects and a country indicator for the United Kingdom are included as in the main specification.

The results are broadly consistent with those from the continuous specification across dictionary-based populist discourse measures (see Figure E.1.1), LLM-based ideology scores, and BERTopic issue shares, though some effects are attenuated as expected given that the binary indicator compresses within-window variation.
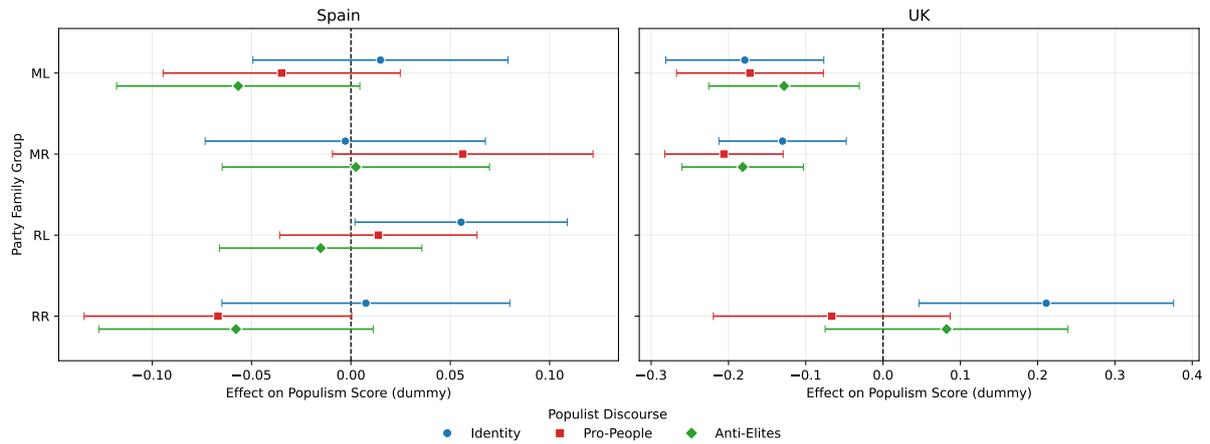
Figure E.1.1: Dictionary-Based Populist Discourse by Election Proximity (dummy), Measure, Party Family, and Country

Notes: Points indicate the average relative transcript-based dictionary score for each party family within each country. The three panels show average levels of identity, anti-elite rhetoric, and pro-people rhetoric. Colors distinguish party families: mainstream left, mainstream right, radical left, and radical right.

Turning to Figure E.1.2 for the BERTopic results, in the United Kingdom, the core findings hold: the mainstream left increases emphasis on environment and services/infrastructure, and the mainstream right on tax/fiscal issues. For the radical right, regional issues remain prominent, and while immigration is no longer significant, defense emerges instead, both of which are issues this party family "owns", consistent with our expectation that outsider parties preserve distinctiveness through selective issue emphasis. In Spain, the mainstream and radical left continue to increase attention to party and ideological framing while reducing emphasis on economy and EU/Brexit, whereas the mainstream and radical right show weaker and less systematic adaptation.
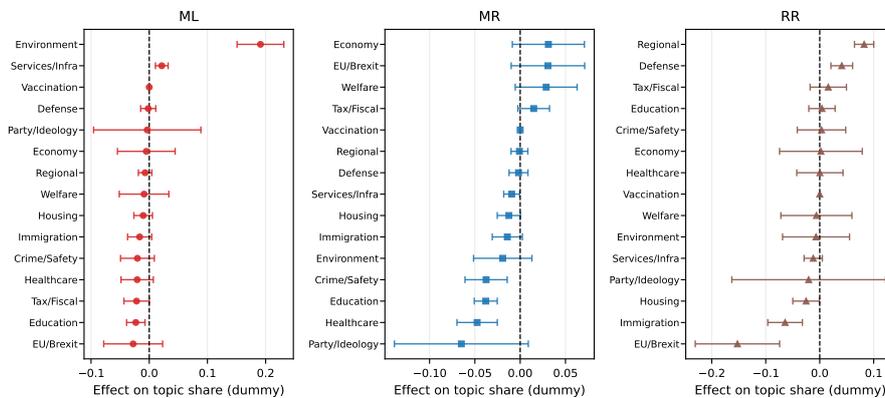
(a) Spain
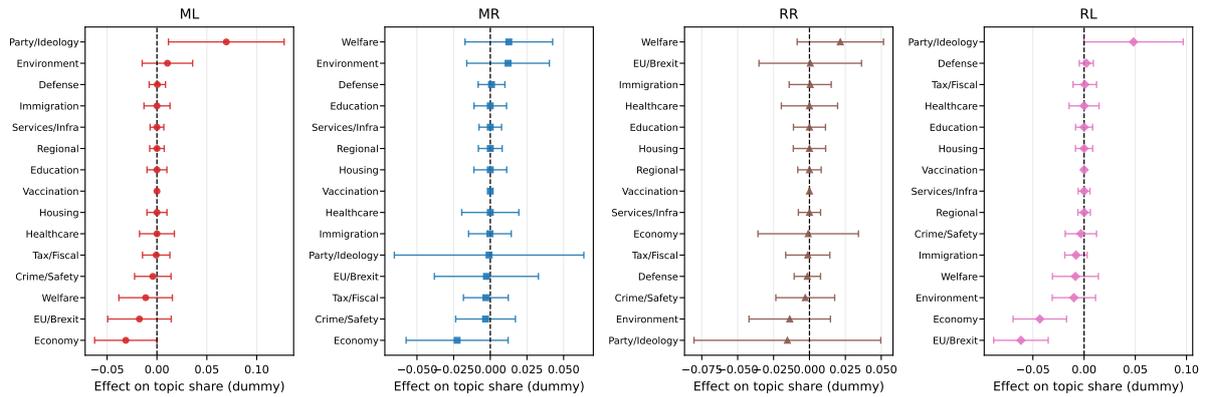
(b) United Kingdom

Figure E.1.2: BERTopic and Election Proximity (dummy) by Country, Party Family, and Issue

Notes: The dependent variables are transcript-based dictionary measures of identity, pro-people rhetoric, and anti-elite rhetoric. The main independent variable is a continuous measure of days closer to election, constructed as the number of calendar days between the video posting date and rescaled so that larger values indicate greater proximity to election day. Estimates come from pooled interaction models with party-family fixed effects, election-period fixed effects, and a country fixed effect for the United Kingdom (Spain is the reference category). The estimation sample contains $N = 4{,}276$ daily family observations. Within each country-family panel, topics are ranked from the largest to the smallest estimated close-election effect.

F ANALYSIS: VAR ANALYSIS

F.1 BERTopic VAR

Figure F.1.1 reports the VAR results for BERTopic issue shares. The VAR estimates 15-day cumulative impulse responses for each topic separately, meaning that cross-party responsiveness can only be estimated for topics on which parties post content within the same temporal window. As shown in Table F.1.1, topic shares are generally low–with the exception of EU Integration & Euroskepticism in Spain (14.53%) and Party Politics & Ideology in the United Kingdom (32.70%), most topics account for less than 10% of daily content in both countries. This sparsity, combined with the requirement that topic attention be sufficiently concentrated over time for the VAR to detect cross-party dynamics, limits the number of estimable responses relative to those shown in Figure 9. Against this backdrop, the results do not reveal clear or systematic patterns of cross-party responsiveness in issue attention, suggesting that parties' topic-by-topic agendas are not strongly coupled at the daily frequency captured by the VAR.
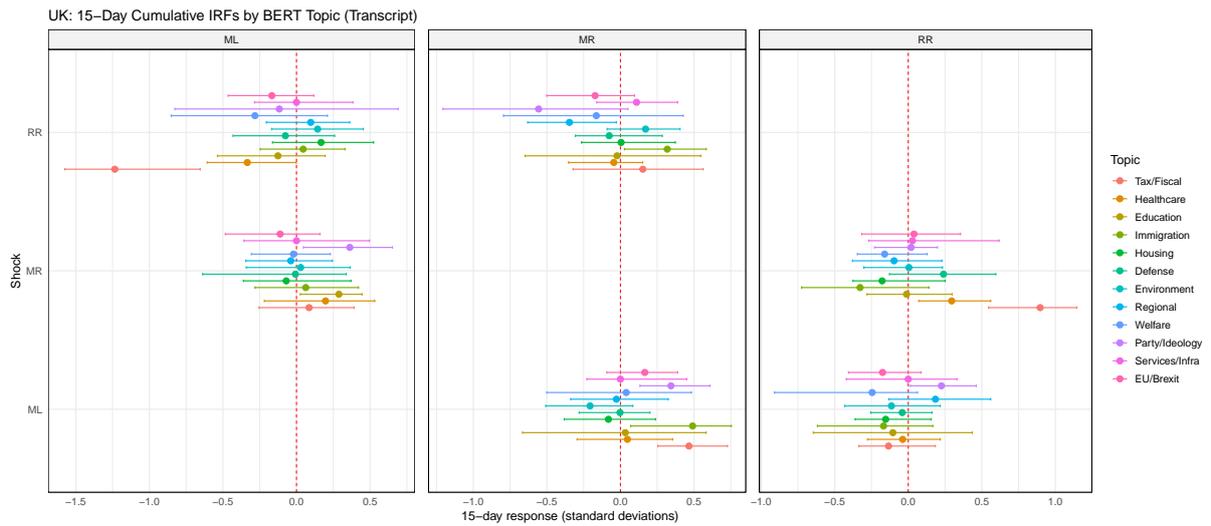
	Spain			United Kingdom	
Rank	Topic	Avg. Share (%)	Rank	Topic	Avg. Share (%)
1	EU Integration & Euroskepticism	14.53	1	Party Politics & Ideology	32.70
2	Environment & Climate Policy	9.69	2	Economy, Business & Employment	6.47
3	Party Politics & Ideology	7.91	3	Social Policy & Welfare	4.41
4	Healthcare System	6.09	4	EU Integration & Euroskepticism	2.80
5	Crime, Policing & Public Safety	4.93	5	Environment & Climate Policy	1.91
6	Immigration, Asylum & Refugees	3.06	6	Crime, Policing & Public Safety	0.72

TABLE F.1.1: Average BERTopic Share per Topic by Country

Notes: Average BERTopic share per topic, expressed as a percentage. Each entry is the grand mean of daily family-level shares, averaged across party families within each country. Topics are ranked from highest to lowest share within each country.
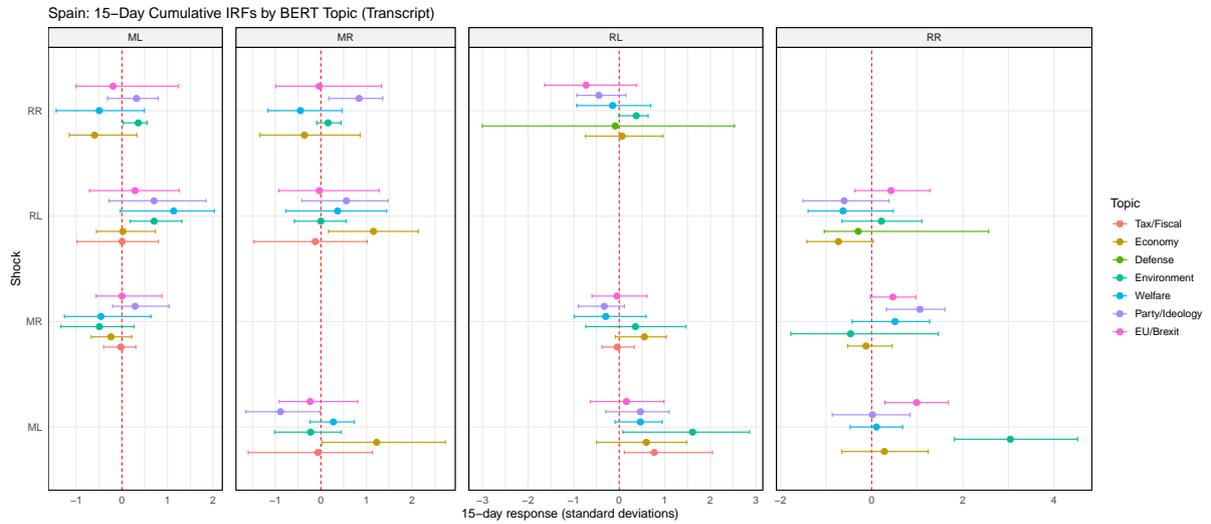
(a) Spain

(b) United Kingdom

Figure F.1.1: VAR Results for BERTopic Responses in Spain and the United Kingdom

Notes: Panel (a) reports Spain and panel (b) the United Kingdom. The dependent variables are transcript-based BERTopic harmonized topic shares. The figures display cross-group 15-day cumulative impulse-response estimates from separate VAR(4) models estimated on transcript-based harmonized BERTopic topic shares. For each topic-specific specification, the dependent variables are the daily topic shares for the party-family series, and the regressors include four lags of each series; election-period dummies are included as exogenous controls. The overlay reports off-diagonal responses only, so own-group responses are omitted. Topics without estimable cross-group intervals are not shown. Points denote standardized effects and horizontal bars indicate bootstrap confidence intervals. The estimation sample includes up to $N_{\text{Spain}} = 1{,}483$ daily observations and $N_{\text{UK}} = 2{,}182$ daily observations.