Strategic Heterogeneity in Policy-Level Positioning: Evidence from Congressional Campaigns*

Colin R. Case† Rachel Porter‡

July 8, 2025

Abstract

To what extent do politicians adopt consistent left-right policy positions, and under what electoral conditions does this consistency vary? These questions are central to the study of representation in American politics, yet current research provides limited insights into how elites position themselves on policy. We address this gap by combining an original corpus of campaign platforms with machine learning methods to estimate the left-right orientations of congressional candidates (n = 4,505) across six salient issue areas in U.S. House elections (2018-2022). Employing our validated measures, we uncover substantively meaningful within-candidate variation in positioning across issues and demonstrate that this variation is systematic. We show that (1) candidates' policy positions closely align with district-level policy preferences, even after accounting for overall district ideology, and (2) policy-level positional inconsistency increases as a candidate's district becomes less favorable for their party. These results challenge prevailing expectations about the extent and sources of heterogeneity in elite policy positioning, underscoring the need for policy-level measures to understand how candidates navigate complex electoral incentives.

Word Count: 9,549

Keywords: elite behavior, congressional campaigns, strategic position taking, ideological constraint, campaign websites, text-as-data

*Authors are listed alphabetically, and both contributed equally. We are thankful for the comments and feedback from Michael Bailey, Daniel Butler, David Fortunato, Anthony Fowler, Jeff Harden, Jaclyn Kaslovsky, Isaac Mehlhaff, Lucia Motolinia, Benjamin Noble, Santiago Olivella, Joe Ornstein, and Julia Park, as well as participants in the 2024 Vision in Methodology Conference. Elizabeth Borowiak, Ryan Casey, Erik Johansen, and Audrey Sachleben provided excellent research assistance. Unite America provided funding for elements of this project.

†Assistant Professor, Department of Political Science, University of Iowa, 359 Schaeffer Hall, Iowa City, IA 52242, colin-case@uiowa.edu.

‡Notre Dame du Lac Assistant Professor, Department of Political Science, University of Notre Dame, 2057 Jenkins Nanovic Halls, Notre Dame, IN 46556, rachel.porter@nd.edu.

Electoral campaigns are central to representative democracies, serving as the primary venue through which candidates communicate their positions on issues to voters, who then select representatives who are aligned with their preferences. Evaluating the quality of democratic representation thus requires an understanding of how candidates choose to position themselves in policy conflicts that define modern politics. Congressional candidates are most often thought to run on cohesive ideological brands, taking consistently extreme or moderate left-right stances on key issues like abortion, immigration, and education. This conceptualization of policy positioning, where elites exhibit a high degree of ideological constraint, underpins a substantial body of research on democratic responsiveness, issue polarization, and electoral accountability (e.g., Ansolabehere et al. 2001; Burden 2004; Bafumi and Herron 2010; Bonica 2014; Hall 2015; Anderson et al. 2024; Case 2025). It also aligns with growing emphasis on the nationalization of American politics, where partisan identity dominates electoral choice (e.g., Lee 2016; Hopkins 2018; Carson et al. 2020), potentially encouraging candidates to prioritize alignment with national party positions over local policy preferences.

Despite the importance of these theoretical foundations, empirical support for consistency in congressional candidates' left-right campaign positioning remains limited. The established expectation of ideological constraint among elites stems primarily from evaluations of roll-call voting behavior within Congress (e.g., Poole and Rosenthal 2011). However, growing scholarship documents considerable heterogeneity in legislators' voting behavior across issue areas (e.g., Lauderdale 2010; Crespin and Rohde 2010; Roberts et al. 2016; Moser et al. 2021). Moreover, congressional campaigns differ fundamentally from legislative contexts: they offer candidates maximum flexibility to tailor their messages without the constraints of party discipline or agenda-setting. Yet existing research that documents policy positioning in U.S. elections focuses on the *types* of issues candidates highlight in their campaigns (e.g., Petrocik et al. 2003; Sulkin 2005; Egan 2013), leaving variation in candidates' left-right stances *across* issues largely unexplored.

Contrary to prevailing assumptions about ideological consistency among congressional candidates, we argue that electoral incentives should motivate meaningful heterogeneity in their mod-

erate to extreme left-right policy positioning. Such strategic behavior is consistent with work demonstrating that voters often hold inconsistent ideological preferences across issues (e.g., Treier and Hillygus 2009; Broockman 2016; Lauderdale et al. 2018; Freeder et al. 2022) and prefer candidates whose policy-level positions closely match their own (e.g., Abbe et al. 2003; Ahler and Broockman 2018; Henderson et al. 2022). If electorates engage in any degree of issue voting, then candidates should, in expectation, tailor their policy-level stances to district-specific preferences on individual issues in order to maximize their own electoral appeal. When district constituencies hold disparate moderate and extreme left-right stances on policy, this heterogeneity should be reflected in office-seekers own issue positions. We expect positional inconsistency to be most pronounced among candidates running in districts that do not favor their own party, where candidates cannot rely solely on co-partisan support and must appeal to cross-pressured voters from the opposing party (Fenno 1978; Burden 2004). In such contexts, adopting disparate left-right positions across issues is a strategic choice for candidates, allowing them to attract voters across partisan subconstituencies—a strategic flexibility that ideological consistency prohibits.

To evaluate strategic heterogeneity in campaign policy-level positioning, we develop novel estimates of U.S. House candidates' left-right positions across six issue areas: abortion, education, energy, guns, healthcare, and immigration. To characterize candidates' policy stances, we draw on an original corpus of campaign platforms scraped from candidates' campaign websites (n = 4,505 or 75% of all major-party, ballot-eligible candidates) for primary elections held between 2018 and 2022. To extract and scale latent policy positions from these texts, we first employ an ensemble of classification models to identify campaign platform paragraphs related to our issue areas of interest. Then, we locally estimate embeddings for words in our corpus (word embeddings) and candidates' issue-specific text for a given election cycle (candidate-issue embeddings). Finally, to compute our positioning estimates, we compare the similarity of candidate-issue embeddings to word embeddings representing the left-most and right-most positions on each issue. Greater similarity to right-most (left-most) word embeddings indicates a more extreme right (left) stance. We refer to our resulting estimates as Candidate Positioning Indexes (CPIs).

We follow best practices in constructing our text-to-measure pipeline, justifying our measurement choices, and validating our scaling procedure (see Park and Montgomery 2023). Illustrating the convergent validity of our estimates, we show that CPIs closely correspond with human judgments of left-right positioning in campaign platform texts. Demonstrating the predictive validity of our estimates, we find that candidates with far-left or far-right issue stances tend to receive PAC contributions from special interest groups that support those corresponding positions.

Our empirical analysis of heterogeneity in policy-level positioning using CPIs proceeds in three stages. First, we explore the extent of substantive within-candidate positional variation across issues. We find modest to weak pairwise correlations in CPIs, suggesting that candidates are not ideologically consistent in their left-right positioning across issues. On average, candidates shift 20 percentile ranks in relative positioning from one issue to another. We demonstrate that positional shifts of this magnitude are substantively meaningful through both qualitative and quantitative examinations of textual similarity. Second, we investigate whether this within-candidate variation in issue-by-issue positioning reflects strategic electoral considerations. In linear regression analyses, we find that district-level, issue-specific public opinion estimates align with CPIs, even after accounting for the overall ideological lean of a district. Building on this evidence for strategic positioning, we then assess the electoral conditions under which candidates engage in weaker ideological constraint, focusing specifically on the dynamics of partisan competition. Employing cross-sectional and two-way fixed effects models, we find that candidates take more variable left-right policy positions as their district becomes less favorable towards their own party.

These findings challenge fundamental assumptions about elite political behavior and carry important implications for democratic representation. By demonstrating that candidates strategically vary their positions across issues rather than maintaining ideological consistency, our results reveal that electoral incentives can promote policy-level representation even in an era of increasing partisan polarization. Identifying when elite issue polarization attenuates and, importantly, the policy domains where moderation is likely to occur is a necessary first step to addressing legislative gridlock and dysfunction (Curry and Lee 2020; Levendusky 2023). Our nuanced view of candidate

positioning suggests that pathways for policy compromise are issue-specific and more likely to emerge among politicians hailing from competitive electoral environments. Moreover, the raw text data and issue-specific measures of policy positioning introduced in this paper constitute a major methodological contribution to the study of elite behavior, electoral institutions, and voting behavior in American politics. Our approach provides scholars with new tools for examining candidate strategy and voter-elite linkages with unprecedented granularity. These resources open numerous avenues for future research into ongoing debates about politicians' convergence on voter preferences, the extent of ideological choice among candidates, the homogeneity of preferences within political parties, and potential avenues for policy compromise in today's polarized Congress.

Elite Policy Positioning: Existing Evidence and Limitations

Scholars of the U.S. Congress and congressional elections have traditionally assumed that politicians exhibit a high degree of ideological constraint—that is, their positions form coherent left-right patterns, where stances on one issue predict stances on all others. Under this assumption, unidimensional ideal point measures for politicians' overall ideology effectively represent policy preferences across diverse issue areas. Ideological consistency among politicians carries significant theoretical and empirical implications, underpinning extensive research within political science and economics that simplifies complex political phenomena with low-dimensional representations (e.g., Poole et al. 1987; Ansolabehere et al. 2001; Clinton et al. 2004; Bonica 2013; Rheault and Cochrane 2020). Some empirical studies of congressional behavior support this assumption: spatial analyses of roll-call voting and interest group ratings indicate that a single liberal-conservative dimension accounts for a large share of variance in legislative vote choice (e.g., Poole and Daniels 1985; Poole and Rosenthal 2000). Most recently, Marble and Tyler (2022) demonstrate high predictive accuracy when applying one-dimensional ideal point models to classify a held-out set of roll-call votes made by federal and state legislators.

However, the roll-call record may provide an incomplete representation of legislators' policy preferences. Party leadership exercises substantial influence over members' voting decisions (Pearson 2015), such that votes on legislation may better reflect party discipline rather than leg-

islators' revealed policy preferences (Snyder and Groseclose 2000; Clinton 2012). Furthermore, many salient issues never receive roll-call votes or appear only within larger omnibus bills that encompass multiple substantive topics (Fowler and Hall 2016; Shin 2024). This selection bias in the roll-call record across issue areas can lead to inaccurate inferences about elite policy preferences. Indeed, Bateman et al. (2017) demonstrate that correcting for policy agenda shifts in Congress across time generates substantively different conclusions about the temporal dynamics of partisan polarization. These findings suggest that roll-call-based positioning measures may overstate the coherence of legislators' stances across issue areas.

Even still, these constrained data reveal some heterogeneity in policy positioning when legislators' votes are disaggregated by issue area. Jochim and Jones (2013) examine the relationship between party unity and issue dimensionality, finding considerable variability in the structure of legislative choice across issue areas. At the individual level, Moser et al. (2021) identify a marked proportion of members who possess "deviant preferences," defined as issue-specific positions that are off-axis from their overall ideological positioning. These findings align with other work demonstrating that members do not maintain ideological consistency in their roll-call voting (e.g., Roberts and Smith 2003; Crespin and Rohde 2010; Lauderdale 2010; Yoshinaka and Grose 2011; Roberts et al. 2016; cf. Fowler and Hall 2016).

If candidates exhibit heterogeneous issue-level positions, traditional unidimensional measures of candidate ideology will obscure critical variations in their actual policy commitments. Such measures will capture an individual legislator's degree of *positional consistency* across policy domains but reveal little about *views on issues themselves* (Broockman 2016). For instance, knowing that a legislator holds liberal positions on two-thirds of issues tells us nothing about which specific issues comprise that majority. Aldrich et al. (2014) raise this point when examining the dimensionality of scaled roll-call votes in the U.S. Congress, concluding that "a one-dimensional dominant result may reflect party 'teamsmanship'...we can tell only that parties are divided from one another, but not if they are divided on one issue, many issues, or...none at all" (p. 438). This limitation constrains our understanding of representation and hampers scholars' ability to examine

how politicians strategically position themselves within specific policy domains—a growing area of research that requires policy-level measurement approaches (e.g., Aldrich et al. 2014; You 2023; Moskowitz et al. 2024; Porter 2022; Porter et al. 2025).

Despite extensive research on legislative policy positioning, scholars have devoted surprisingly little attention to assessing politicians' positioning in campaigns, particularly the degree to which candidates maintain consistent left-right positions across issues. Numerous studies develop approaches to estimate non-incumbent candidates' overall political orientations, employing data on candidate survey responses (e.g., Ansolabehere et al. 2001; Montagnes and Rogowski 2015), social media behavior (e.g., Barbera 2015; Macdonald et al. 2025), fundraising networks (e.g., Bonica 2014; Hall 2015), and campaign website content (e.g., Pennec et al. 2024; Bailey 2024; Case 2025). Yet these measurement strategies share the same critical limitation described above: they compress candidate positioning into a single dimension, assuming away the possibility of ideological heterogeneity across policy domains. Consequently, some scholarship (e.g., Tausanovitch and Warshaw 2017) raises fundamental questions about these measures and their capacity to accurately represent candidates' substantive policy orientations.

While some scholars have examined issue-specific positioning in congressional campaigns, this work primarily focuses on issue ownership (e.g., Petrocik et al. 2003; Egan 2013; Hayes and Lawless 2016) and issue uptake (e.g., Kaplan et al. 2006; Sulkin 2005, 2011), questions that address the dynamics of issue *selection* rather than ideological *stance*. In exception, Moskowitz et al. (2024) produce estimates for candidates' left-right policy-level positions for U.S. House races from 1996 to 2008 using Project Vote Smart's survey of federal candidates (NPAT). The authors find that cultural, rather than economic, issues drive increasing polarization via member replacement. Bonica and Li (2021) employ a supervised machine learning approach to infer candidates' issue positions based on the itemized contributions they receive, finding a high degree of correspondence between campaign issue scores and interest group issue ratings within Congress. Yet these studies do not address the prevalence of positional constraint in congressional campaigns and the strategic con-

siderations that lead candidates to maintain or deviate from ideological consistency.[1]

This accumulated evidence highlights the empirical challenges associated with measuring the policy preferences of politicians. As Lee (2015) notes, "Aggregating across the wide and changing range of issues facing the federal government to arrive at a single summary measure...raises thorny methodological questions" (p. 267). Beyond these measurement difficulties, significant conceptual gaps persist in our understanding of policy-level positioning. Although extensive literature debates the presence and extent of ideological constraint in legislative behavior, this debate has not been extended to congressional campaigns, where candidates face fundamentally different institutional constraints and strategic incentives than those governing legislative decision-making.

A Theory for Strategic Heterogeneity in Policy-Level Positioning

Should congressional candidates maintain consistent left-right positions across issues, or strategically vary their stances? There are compelling reasons for candidates to employ ideological constraint. Adopting consistent policy stances helps candidates build clear ideological brands, reducing voter uncertainty about their preferences (Downs 1957; Kingdon 1973; Snyder and Ting 2002). This clarity in positioning can have electoral consequences given that voters dislike ambiguity (e.g., Alvarez 1997, cf. Tomz and Houweling 2009) and prefer ideological consistency (e.g., Rogowski and Tucker 2018, cf. Somer-Topcu 2015). Additionally, positional inconsistency might invite attacks from opponents (Fenno 1978), increasing the electoral costs of such inconsistency. However, as Downs (1957) notes, this strategy entails a tradeoff: while a uniform policy orientation may enhance appeal to ideologically aligned voters, selectively adopting positions across the left-right spectrum may broaden a candidate's appeal.

Resolving this tension requires understanding how citizens organize their policy preferences across different issue domains. As single-minded election seekers (Mayhew 1974), candidates should tailor their campaign messages to increase their chances of winning office. If voters engage

[1] Rogowski and Tucker (2018) measure ideological consistency among congressional candidates using data from a survey of expert informants in 131 U.S. House races in 2006. The authors define candidates' ideological placement as the experts' average candidate rating, and variability in positioning as the variance in experts' ratings. This approach differs from our own in that we measure candidate positioning and variability using candidates' expressed policy stances rather than perception-based data.

in issue voting, candidates should strategically align their policy positions with district-specific preferences to maximize electoral prospects. When constituencies have diverse left-right preferences across issues, we posit that candidates' own issue positions should mirror this diversity. According to Downs (1957), positional inconsistency is most advantageous in electoral contexts where candidates need to generate broad appeal. Therefore, we anticipate this behavior to be most evident in districts where candidates cannot rely solely on same-party voter support and must attempt to cross-pressure out-party or swing voters.

The Nature of Voter Policy Preferences

Converse (1964) introduced the term "constraint" to describe the degree to which Americans' opinions on various issues are correlated, concluding that a single liberal-conservative dimension fails to structure most Americans' views across the full range of policy issues. Contemporary research reinforces this finding by demonstrating that unidimensional measures explain only 20-40% of the variation in expressed policy opinions among voters (Lauderdale et al. 2018; Freeder et al. 2022; Marble and Tyler 2022; Broockman and Lauderdale 2025). This variation in citizens' preferences manifests in systematic patterns: Broockman (2016) reveals stronger correlations in citizens' preferences within issue areas than across them, while Warshaw and Rodden (2012) documents substantial variability in issue-specific public opinion both within and across congressional districts. These findings collectively suggest that voters' left-right positions vary meaningfully across different policy domains rather than adhering to a consistent ideological orientation.

Heterogeneity in mass preferences carries important implications for voting behavior. Although citizens value multiple dimensions of representation (see Harden 2015), they especially prioritize policy representation on salient issues of the day (Lapinski et al. 2016) and personally important issues (Ryan and Ehlinger 2023). These policy preferences influence vote choice. Henderson et al. (2022) show in an experimental setting that voters consistently and effectively use policy information when evaluating candidates. In an original survey of primary and general election voters, Colao et al. (2025) show that respondents were substantially more likely to vote for a candidate with whom they have policy agreement—even if that candidate is an out-partisan. Most

relevant for our purposes, Ahler and Broockman (2018) demonstrates that citizens prefer candidates who align with them on individual issues rather than candidates whose overall ideological profile is closer to their own, providing direct evidence that issue-specific congruence matters more than broad ideological similarity for vote choice.

Strategic Implications for Candidate Policy-Level Positioning

If voters possess diverse left-right policy opinions and engage in issue voting, this creates a strong incentive for office-seekers to align their own positioning with these varying preferences. We argue that this strategic adaptation will manifest as within-candidate variation in left-right positioning across different policy domains. When voter preferences are distributed heterogeneously across issue spaces, optimal candidate positioning requires issue-specific adjustments rather than adherence to a single ideological location (Downs 1957). This expectation aligns with extensive research indicating that a host of factors influence a candidate's overall positioning, including their electoral context (e.g., Burden 2004; Case 2025) and personal valence qualities (e.g., Stone and Simas 2010). Given that candidates tailor their overall ideological orientations, we posit that they should exhibit similarly strategic behavior in issue-specific positioning.

The logic of issue-specific positioning becomes particularly complex when considering the electoral calculus faced by candidates running in districts where both major parties have a viable chance of winning the seat or the other party is favored. In these contests, candidates face a strategic dilemma: they must appeal to their primary constituency enough to secure the nomination, but still maintain electability in the general election (Brady et al. 2007; Anderson et al. 2024). Candidates are also limited in the extent to which they can shift their issue positions between the primary and general election, facing constraints from their own beliefs, the difficulties of communicating a change in position to voters, and the disadvantages of appearing indecisive or opportunistic (Powell 1982). Rather than attempting to appeal to diverse subconstituencies through a single moderate ideological signal, which may dissatisfy both in-party and out-party supporters, we posit that candidates running in districts where their party does not have an electoral advantage will attempt to build cross-party coalitions by adopting disparate left-right positions across multiple policy do-

mains. Conversely, in same-party strongholds, candidates tailor their positioning to the preferences of primary voters, who constitute the most consequential electoral gatekeepers. Adopting greater ideological constraint in this context may be more strategic, given that primary voters tend to be more politically knowledgeable, ideologically extreme, and exhibit a higher degree of ideological constraint themselves (Brady et al. 2007; Hill 2015; Barber and Pope 2018).

In the sections below, we outline our methodology for scaling the left-right policy-level positions of office-seekers. After validating these measures, we explore strategic positioning dynamics by examining the central expectations outlined above. Our findings provide novel insights into the relationship between district-level electoral competition, issue representation, and candidate positioning across issue domains.

Measurement Approach

We measure policy positioning among congressional candidates with an original collection of campaign platforms from the websites of all available U.S. House candidates in 2018, 2020, and 2022. We produce Candidate Positioning Indexes (CPIs) for six issue areas: abortion, education, energy, guns, healthcare, and immigration—issues identified by Pew Research as top public priorities and "very important" to vote choice.[2] Using this corpus, our text-to-measure pipeline proceeds in three stages: First, we identify all paragraphs from each candidate's platform relating to our six issues of interest. Second, we estimate embeddings for every word in the vocabulary of candidates' campaign platforms and all issue-relevant text in each candidate's campaign platform for a given election cycle. Third, we measure the similarity between these candidate-issue embeddings and the averaged word embeddings for dictionaries of terms indicating the left-most and right-most positions on each issue. Resulting similarity measures constitute our CPIs.

Data: Online Campaign Platforms

Per Druckman et al. (2009), an ideal data source on policy positioning in campaigns will be "unmediated (i.e., directly from the campaign), complete (i.e., covering a full range of rhetorical

[2]These issues were identified by the Pew Research Center as top policy priorities among the general public and "very important" to vote choice among registered voters in election years of interest.

strategies), and representative of the population of campaigns" (p. 345). However, existing data sources fail to meet these criteria. Candidate surveys (e.g., National Political Awareness Test) directly measure comprehensive policy stances but suffer from low response rates and exclude primary candidates—a significant limitation as primaries have become the main venue for meaningful congressional competition (Jacobson and Carson 2016). Other measures achieve greater coverage using data on perceptions of candidates' positions rather than their stated stances (e.g., Stone and Simas 2010; Barbera 2015; Bonica 2014. However, perception-based data may be influenced by non-policy factors (Tausanovitch and Warshaw 2017) and rarely include policy-specific evaluations necessary for scaling candidates' positions on individual issues.

We employ an original corpus of campaign websites to address these limitations. Per Porter et al. (2024), campaign websites are ubiquitous in modern campaigns. Websites feature candidates' policy-specific positions in their own words across numerous issue areas, with no time or space restrictions, allowing discussion of every issue important to their campaign (Sulkin et al. 2007). Candidates and their staff spend substantial time crafting their website messaging because these sites serve as an informational "hub" in campaigns (Druckman et al. 2009). For this reason, over a dozen states include links to campaign websites on official listings of ballot-eligible candidates.[3]

To compile our original corpus of campaign platforms, we first generated a list of all candidates running for the U.S. House in a given year and then located the URL for each candidate's campaign website, if available. Of the 6,006 ballot-eligible, major-party candidates who ran between 2018 and 2022, we identified 5,229 (87%) as having a campaign website. We characterize campaign platforms as the text presented on the "Issues" or "Policy Positions" page of these campaign websites. Teams of research assistants were tasked with locating and extracting platform text from websites via manual downloading. Website text was collected less than two weeks before a candidate's primary election to maximize data coverage and consistency. We document our data collection strategy in greater detail in Appendix A1.[4] We identified 4,505 candidates as having a

[3]As of 2024, these states include Alaska, Arizona, Delaware, Georgia, Hawaii, Illinois, Maryland, Minnesota, Montana, New Hampshire, Oregon, Vermont, and Virginia.

[4]We are not the first to collect text data from websites (e.g., Xenos and Foot 2005; Dolan 2005; Sulkin et al. 2007; Druckman et al. 2009; Bailey 2024; Meisels 2024; Pennec et al. 2024). However, our data collection effort is

website-hosted campaign platform (75% of major party candidates). Candidates without campaign platforms or, more broadly, campaign websites tend to be non-competitive challengers.[5]

Method: Semantic Projection

Isolating Issue-Specific Text

To produce CPIs for our six issues of interest, we rely exclusively on issue-specific text. Restricting the breadth of text used in our scaling procedure helps to limit noise in measurement (Grimmer and Stewart 2013; Egerod and Klemmensen 2020).[6] To isolate text, we first spliced campaign platforms into individual paragraphs, producing a corpus of 136,368 natural paragraphs.[7] Next, we randomly sampled 6% of these paragraphs (8,306 total) and labeled them for the presence/absence of content related to our six issue areas of interest.[8] Categorization was not mutually exclusive; multiple issue areas were sometimes discussed within the same paragraph. Full coding instructions can be found in Appendix A2.2.

These labeled paragraphs serve as training data for a series of supervised machine-learning classification models. We trained five classification models for each issue area and employed their predictions to train a stacking ensemble classifier. This approach enabled us to leverage a diverse range of modeling approaches to achieve better performance, taking advantage of the strengths of each model for final classification. Appendix A2.3 provides greater detail on model training and classifier validation. Across all issue areas, our ensemble classifiers achieved an F1 score of at least 0.95 on model training data. When classifying a held-out set of paragraphs from campaign platforms, we achieved F1 scores of 0.82 or higher across all models.

distinct in important ways. First, our population of interest constitutes *all* major-party, ballot-eligible candidates who ran in primaries or general elections. Other collections focus on collecting text for only a subset of all candidates. Second, we collected text in real-time, rather than relying on Internet archives via the Way Back Machine or Library of Congress, which can have idiosyncratic coverage.

[5]See Table A1 for a full analysis of campaign website missingness.

[6]In Appendix A2.1, we demonstrate text inclusion has downstream consequences on CPI estimates.

[7]The choice of aggregation at the natural paragraph level follows best practices from Daubler et al. (2012) that emphasize the importance of exogenous unitization of text. We aggregate short paragraphs (i.e., less than fifteen words) into the most proximate multi-sentence paragraph.

[8]Three trained readers labeled paragraphs for policy content; two coders labeled 3,000 paragraphs, and one coder labeled 2,306. Of the 8,306 paragraphs, all three coders labeled 308 (4%) to assess inter-coder reliability. For this reliability set, we achieve 88% agreement across all three readers.

Estimating Word Embedding Model

We next employ a word embedding model to uncover meaning in these issue-relevant paragraphs. Word embedding models use a neural network architecture to predict word(s) in a document given the word(s) that occur in close context to that word. Resulting embeddings are vector representations of semantic relationships between words in a dense, continuous, high-dimensional space. Our estimation approach relies on a Doc2Vec word embedding model. This approach is similar to a traditional word embedding model but allows for the incorporation of covariates in the training process, which produces *covariate-level embeddings* (see Le and Mikolov 2014). In our specification, each paragraph is assigned a covariate that corresponds to the candidate, issue area, and election year.[9] Resulting covariate embeddings—which we refer to as candidate-issue embeddings[10]—are numerical representations constituting a semantic summary of the rhetorical meaning for *all issue-specific paragraphs* from a given candidate in a given year.

Our Doc2Vec model operates in two parts: the first is similar to a traditional skip-gram Word2vec architecture (Mikolov et al. 2013). A word, w_k, is sampled at each iteration, where k is the word's position in a document. A window of length Δ is extracted twice, once before and after w_k. The words in the window surrounding w_k are our outcomes of interest. These outcomes can be written as $w_\Delta = (w_{k-\Delta},...,w_{k-1},w_{k+1},...,w_{k+\Delta})$. The model input is an indicator vector, x_k, for the target word, w_k, which selects the corresponding word embedding, β_k. This embedding predicts each word, m, in w_Δ. Figure 1 illustrates this word embedding model architecture.

The second part of our model uses the same architecture: a word, w_k, is sampled at each iteration, and a window of length Δ is extracted before and after. Words falling in this window, w_Δ, represent our outcomes of interest. Instead of using an indicator vector to index the word embedding for target word w_k, the indicator vector, $z_{i,j,t}$, is used to index the candidate-issue embedding, $\zeta_{i,j,t}$, for candidate i on policy area j in election year t. This embedding predicts each

[9]Recall that we classify some paragraphs (4% of all paragraphs) as discussing more than one issue (e.g., a paragraph discussing abortion and healthcare). In these cases, we include duplicate paragraphs in model training.

[10]A more accurate descriptor would be "candidate-issue-year" embedding, as our approach produces unique embeddings for candidates who discuss the same issue across multiple elections. We use "candidate-issue" for brevity.

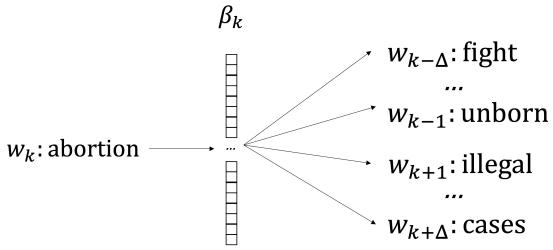


Figure 1: Word Embedding Model Architecture

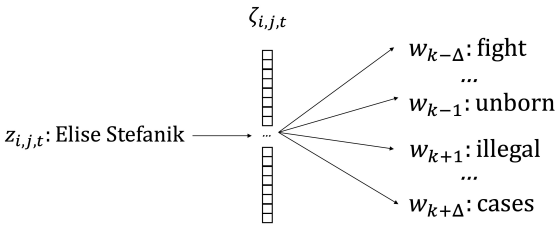


Figure 2: Candidate-Issue Embedding Model Architecture

word, m, in w_Δ.[11] Figure 2 illustrates this candidate-issue embedding model architecture.

We begin with embeddings pre-trained on the Google News corpus to set initial model weights for word embeddings; initial weights for candidate-issue embeddings are random. Our Doc2Vec model alternates between the steps described above during training to tune the parameters within embeddings. We fit our model using parameter recommendations from Rodriguez and Spirling (2022), including a window of six and an embedding dimension of 300. We employ default hyper-parameters from the original Doc2Vec model. We take standard text pre-processing steps for embedding models, as laid out by Rodriguez and Spirling (2022).[12]

Our model produces two sets of outputs. First, our model produces 11,577 word embeddings (β_k)—an embedding for every word in our corpus. Second, our model produces 18,630 covariate embeddings ($\zeta_{i,j,t}$)—an embedding for every candidate-issue-year combination in our data. If our classifier identified no issue-relevant paragraphs for a given candidate in a given year, then we do not generate an embedding for that candidate-issue-year combination. Our modeling approach places covariate and word embeddings in the same dimensional space; we leverage this feature to produce our CPI measures in the following section. Additionally, our model is trained on texts generated for all included election years, facilitating cross-cycle comparisons of policy positioning.

[11]This process is intuitively similar to taking the average of the word embeddings for candidate i on policy area j in year t. However, the Doc2Vec estimation places less weight on high-frequency words (e.g., the, is, in) that often possess little rhetorical meaning. For this reason, previous work finds Doc2Vec outperforms averaging word embeddings (Lau and Baldwin 2016; Grimmer et al. 2022).

[12]These pre-processing steps minimally alter the underlying meaning of text (Denny and Spirling 2018). However, in several important instances, these steps *do* affect the underlying meaning of words in our corpus. For this reason, we modify text in several instances to preserve underlying meaning (e.g., converting instances of "2nd amendment" to "second amendment"). The full list of our steps in text prep-processing is available in Appendix A2.4.

Table 1: Issue Cleavages & Positional Endpoint Term Dictionaries

Issue Area	Left-Most & Right-Most Term Dictionaries
Abortion	*Pro-Choice* (L): reproductive, justice, prochoice, codifying, autonomy, personal... *Pro-Life* (R): birth, unborn, conception, heartbeat, prolife, ban, born, outlaw...
Education	*Federal* (L): public, invest, free, tuition, universal, prek, equality, salaries, teachers... *Local* (R): parents, local, control, decentralize, choice, competition, charter, private...
Energy	*Renewables* (L): renewable, net, zero, incentives, credits, subsidize, develop... *Fossil Fuels* (R): oil, coal, deregulate, repeal, free, market, domestic, reserves...
Guns	*Restrictions* (L): mandatory, background, ban, national, registry, assault, automatic... *Access* (R): infringed, inherent, unrestricted, repeal, reciprocity, concealed, carry...
Healthcare	*Publicly Funded* (L): universal, medicareforall, singlepayer, human, right... *Privatized* (R): free, market, open, competition, choice, consumers, deregulate...
Immigration	*Inclusive* (L): undocumented, daca, pathway, dreamers, dignity, abolish, ice... *Exclusive* (R): enforce, law, build, wall, verify, sanctuary, chain, birthright...

Scaling Candidate Policy Positioning

In this final stage, we scale candidates' policy positioning on a left-right dimension that defines discourse for our six issue areas of interest. Issues are multidimensional, and our measures do not capture positioning on all these dimensions; instead, we pre-specify a dominant cleavage on which to scale each issue area. Italicized text in Table 1 denotes cleavages of interest.

To produce our issue-specific measures, we compare the relative closeness of each candidate's covariate embedding, $\zeta_{i,j,t}$, to the left-most and right-most positions for a given issue cleavage. We define these positional endpoints using dictionaries of terms and their associated word embeddings. Using dictionaries to capture key quantities of interest is standard in research employing word embeddings (e.g., Kitagawa and Shen-Bayh 2024; Garg et al. 2018; Garten et al. 2018). We selected terms based on the close reading of policy platforms from far-left and far-right advocacy groups. In Appendix A2.5, we provide greater detail on our term selection and demonstrate that our estimated CPIs are robust to alternative term selections. Table 1 outlines an abbreviated list of terms in each left/right term dictionary; Appendix Table A11 presents the full dictionaries.

The full scaling procedure for constructing CPIs is as follows. First, for issue j, we average the embeddings for the set of words, R_j and L_j, in our right and left issue dictionaries, respectively.

Resulting averages produce two embeddings—ρ_j, and λ_j—indicative of the semantic meaning of the right-most and left-most positions for the cleavage defining j. Second, we subtract λ_j from ρ_j to create an axis of left-right positioning, α_j. Extant literature establishes that applying vector algebra to word embeddings produces a semantic axis of meaning (e.g., Bolukbasi et al. 2016; Garten et al. 2018; Kozlowski et al. 2019; Grand et al. 2022). In our application, this axis represents the left-right dimension for the cleavage that defines j. Finally, for each candidate i in year t with an estimated covariate embedding for issue j, we calculate the cosine similarity between $\zeta_{i,j,t}$ and α_j. This estimation procedure for calculating our Candidate Position Indexes is written as:

$$
\begin{aligned}
R_j &= \{w_1, w_2 ... w_n\} \\
L_j &= \{w_1, w_2 ... w_n\} \\
\rho_j &= \overline{\beta_{k \in R_j}} \\
\lambda_j &= \overline{\beta_{k \in L_j}} \\
\alpha_j &= \rho_j - \lambda_j \\
CPI_{i,j,t} &= cos(\alpha_j, \zeta_{i,j,t})
\end{aligned}
$$

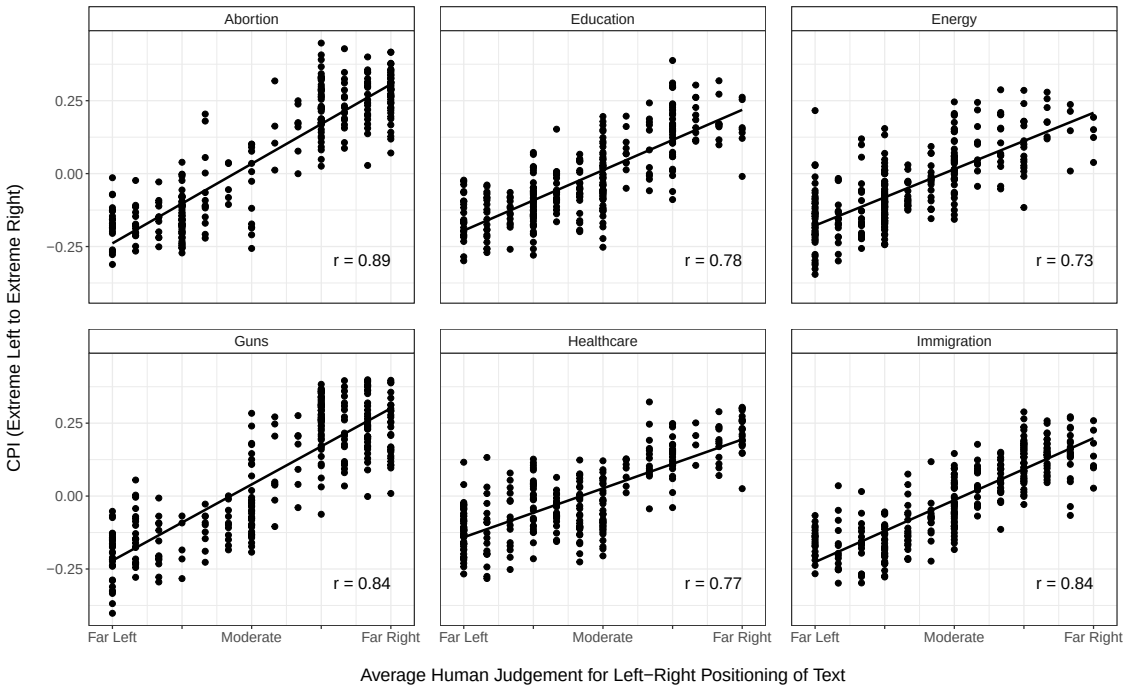
The resulting estimates constitute our Candidate Positioning Indexes (CPI). We standardize CPI with a mean of 0 for ease of interpretation. As CPI increases (decreases), a candidate's text is closer in meaning to the right-most (left-most) position for that issue cleavage.

Measurement Validation

We compare human judgments of policy platform text to our Candidate Positioning Indexes as a benchmark for validity. Per Lowe and Benoit (2013), "valid positional estimates from quantitative scaling, for a given dimension, should match a human reader's placement of these texts with respect to identifying relative differences along this dimension" (p. 300). Given humans' superior ability to read and understand the meaning of natural language, directly comparing the quantitative scaling of text to qualitative human judgments offers a meaningful benchmark for semantic validity.

In this validation exercise, two student readers and one principal investigator (PI) scored the same random selection of documents (n = 900). Texts were provided to readers in their raw form at

Figure 3: CPI Validation with Human Judgments



units of aggregation identical to those employed to generate candidate-issue embeddings. No candidate or party-identifying information was provided beyond what was available in the documents. We provided readers with detailed instructions outlining each issue's cleavage, which included broad guidelines about scoring and policy-specific examples.

Readers scored documents on a five-point integer scale from very left (–2) to very right (+2). When applicable, readers flagged document content as ambiguous or irrelevant to their scaling task. If cleavages of interest were not clear to readers, this would suggest that differences identified through scaling do not match our intended quantities of interest. Across issues, on average, less than 10% of human-coded texts were flagged by any individual reader as irrelevant or ambiguous. Full documentation of our human judgment task, including coder instructions, descriptive statistics on reader performance, as well as sample texts, is provided in Appendix A3.

Figure 3 presents a series of scatter plots comparing human judgments of text with their associated CPI.[13] Each facet corresponds with our six issue areas; plots disaggregated by candidate party are available in Appendix Figure A2. Correlations between human judgments of text and

[13]Text flagged as "ambiguous" or "irrelevant" is set to zero (i.e., centrist).

estimated CPI range from 0.73 to 0.89. These correlations are comparable to those from other studies evaluating the convergent validity of automated scaling methods against human judgments (see Park and Montgomery 2023 and Ruedin and Morales 2019). Results from this validation task are suggestive that CPIs (1) reflect cleavages of interest, as specified in Table 1, and (2) correspond closely with human placements of text on a left-right spectrum.

As a test of predictive validity, we examine the relationship between Political Action Committee (PAC) contributions and CPI extremity. Extensive research documents the giving behavior of groups, finding that these entities contribute to candidates for access-oriented and ideological reasons (e.g., Fouirnaies and Hall 2014; Barber 2016; Meisels 2024). Politicians may even adjust their position-taking behavior to enhance PAC appeal (Baker 2016). For our purposes, we remain neutral regarding the direction of the causal relationship between candidate positioning and PAC contributions. The foundation of our predictive validation exercise is that candidates whose policy positions align more closely with the priorities of PAC-funded groups should receive fundraising contributions from these groups more frequently (Li 2018; Bonica and Li 2021).

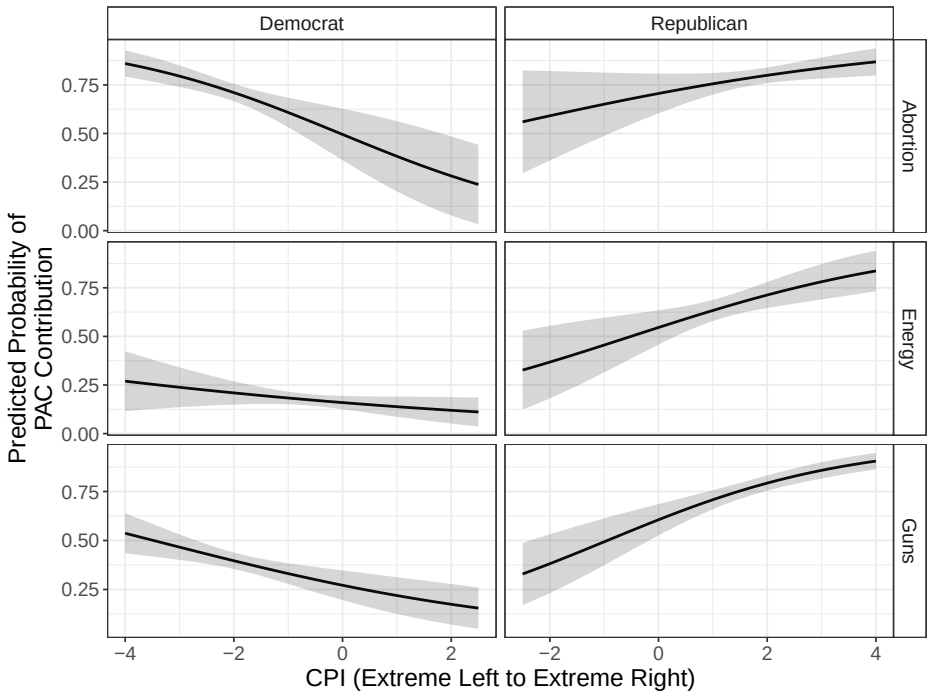
Figure 4 presents predicted probabilities for PAC giving as a function of CPI extremity. The dependent variables across the logistic regressions are binary indicators for the presence (or absence) of a PAC donation from an abortion-related PAC (top facet), an energy-related PAC (middle facet), or a gun-related PAC (bottom facet).[14] We identified PACs as aligned with left (right) leaning positions on abortion, energy, and guns using data from OpenSecrets. We constrain units of analysis to general election contenders, as PAC giving is relatively rare in primaries.[15] We include fundraising data for all giving in support of a candidate (i.e., direct donations and independent expenditures). Models control for unidimensional positioning, as estimated by Bonica (2024).[16] Full model outputs are available in Appendix Tables A12 to A17. Across five of the six panels in Figure 4, we find a statistically significant association between PAC giving and the extremity

[14]In our abortion and gun analyses, PACs constitute ideological/single-issue groups advocating for policy-specific positions. In our energy analyses, PACs constitute private companies in the energy sector.

[15]Units include only candidates for whom we estimate CPI-Abortion, CPI-Energy, or CPI-Guns; as discussed earlier, we do not generate issue-specific CPIs for candidates with no issue-specific text. We replicate this analysis with imputed scores in Appendix Tables A12 to A17. These results are similar to those presented in Figure 4.

[16]Bivariate models produce substantively identical results and can be found in Appendix Tables A12 to A17.

Figure 4: CPI Validation with Position-Aligned PAC Giving



Note: Predicted probabilities for presence/absence of abortion-related PAC (top facet), energy-related PAC (middle facet), or gun-related PAC contribution as a function of CPI-Abortion, CPI-Energy or CPI-Guns. Full model outputs are available in the left-most column of Appendix Tables A12 to A17.

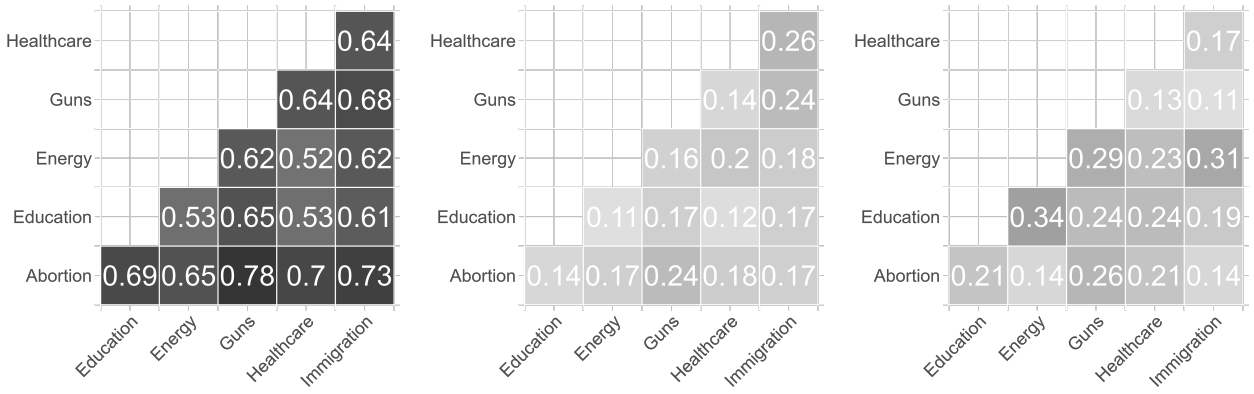
of issue-specific candidate positioning.[17] Recall that negative (positive) CPI values indicate more extreme left (right) policy positioning. For our models of Democratic and Republican candidates, the association between PAC giving and positional extremity is in the expected direction.

Results

The empirical analyses that follow provide support for our three main theoretical expectations. First, candidates exhibit heterogeneous left-right positioning, as demonstrated by modest correlations between issue-specific CPIs and meaningful shifts in their positional extremity across policy areas. Second, this variation appears strategic: candidates' positions consistently mirror constituency preferences, with district-level, issue-specific public opinion predicting candidate positioning even after accounting for overall district ideology. Third, strategic inconsistency increases as a candidate's district becomes less favorable towards her own party.

[17]The relationship between energy PAC contributions and CPI is not significant for primary election winners.

Figure 5: Percentile-Rank Pairwise CPI Correlations, By Candidate Type



Establishing Heterogeneity in Policy-Level Positioning

We start by examining the extent to which congressional candidates adopt consistent left-right policy positions across different issue domains. Figure 5 presents correlation matrices that offer initial descriptive evidence. Each facet displays average pairwise correlation coefficients for CPIs, where raw estimates are converted to percentile ranks to improve interpretability by reflecting relative positions rather than differences in distributions across issue-specific measures.[18] We create percentile rankings by first identifying candidates who discuss both issues in each pairing and then generating separate percentile scores for each issue within this selected group.

The results show limited consistency in candidates' left-right positioning across issues. In the full sample of congressional candidates (left panel, Figure 5), pairwise correlations are generally moderate, with most falling below 0.70. This pattern becomes more pronounced when disaggregated by party: within-candidate correlations across issues fall below 0.40 for both Democrats and Republicans (middle and right panels, Figure 5). These findings suggest that candidates do not present a unified ideological orientation across policy areas, but instead display significant issue-specific variation in their left-right positioning.

To complement our correlational analysis, we measure the magnitude of within-candidate variation in policy positioning across issue pairs. For each candidate who discusses both issues in a

[18]Figure A3 depicts identical takeaways for average pairwise correlation coefficients with raw CPI scores.

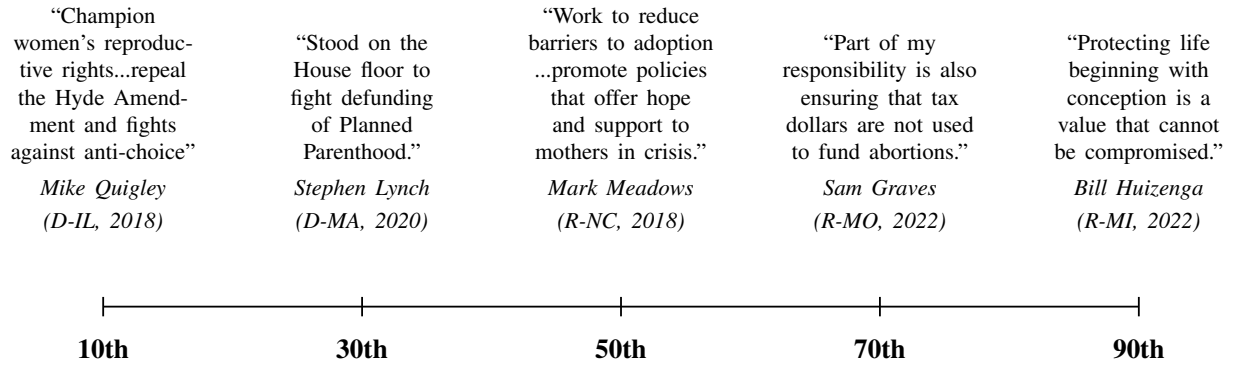
Table 2: Average Absolute Differences in Percentile Rankings for Candidate-Level CPI Pairings

	Abortion	Education	Energy	Guns	Healthcare	Immigration
Abortion	0	15.7 [6.4, 22.8]	18.5 [7.2, 26.7]	20.0 [8.2, 28.7]	17.9 [7.3, 25.9]	17.3 [6.5, 25.6]
Education	18.5 [7.2, 26.7]	0	22.5 [7.8, 33.5]	19.1 [7.2, 27.6]	22.5 [7.6, 33.3]	20.1 [8.0, 28.9]
Energy	20.0 [8.2, 28.7]	22.5 [7.8, 33.5]	0	20.0 [7.1, 28.9]	22.6 [8.1, 34.6]	19.9 [7.2, 28.9]
Guns	15.7 [6.4, 22.8]	19.1 [7.2, 27.6]	20.0 [7.1, 28.9]	0	19.0 [6.9, 26.8]	18.4 [6.9, 26.5]
Healthcare	17.9 [7.3, 25.9]	22.5 [7.6, 33.3]	22.6 [8.1, 34.6]	19.0 [6.9, 26.8]	0	18.9 [6.9, 27.2]
Immigration	17.3 [6.5, 25.6]	20.1 [8.0, 28.9]	19.9 [7.2, 28.9]	18.4 [6.9, 26.5]	18.9 [6.9, 27.2]	0

given pairing, we calculate the absolute difference in their CPI percentile rankings and then average these values across all such candidates to measure the typical size of cross-issue positional shifts. This approach directly quantifies how much candidates' relative positioning shifts between policy domains. Table 2 presents magnitude differences in percentile rankings, with each cell reporting the mean absolute difference in percentile rankings for candidate pairwise CPI comparisons and interquartile ranges in brackets. The results reveal significant positional variation, with candidates shifting their relative positions by about 20 percentile ranks from issue to issue. The interquartile ranges also display notable dispersion, with candidates shifting their relative positioning between 8 and 30 percentile points.

To better contextualize the substantive magnitude of these positional rank shifts, we provide illustrative campaign platform excerpts at 20 percentile-rank intervals for CPI-Abortion. Figure 6 demonstrates that these percentile-point shifts correspond to meaningful differences in candidate position-taking, with stances ranging from strong pro-choice advocacy at the 10th percentile (left) to explicit pro-life positioning at the 90th percentile (right). The substantive progression is clear in specific candidate comparisons: Mark Meadows (R-NC, 50th percentile) emphasizes "reducing barriers to adoption," representing a moderate position focused on alternative solutions to abortion, while Sam Graves (R-MO, 70th percentile) adopts a more conservative stance by focusing on

Figure 6: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Abortion



ensuring "tax dollars are not used to fund abortions." This contrast illustrates how candidates' positions shift noticeably across a 20 percentile-rank interval, with distinct emphases emerging at different points along the left-right spectrum. Complete textual comparisons across all CPI issue areas are available in Appendix Tables A18 to A23.

We systematically assess the magnitude of positional shifts in Table 3 through a series of keyness analyses comparing abortion-related campaign platform statements across adjacent percentile bands for CPI-Abortion.[19] This corpus-linguistic method identifies terms that occur with significantly higher frequency in one set of texts compared to another. Here, we compare adjacent percentile groupings of issue texts by CPI-Abortion (90th vs. 70th, 70th vs. 50th, etc.) to identify words most distinctively associated with candidates in higher (right-leaning) versus lower (left-leaning) percentile groups. This approach isolates the incremental textual shifts that occur as candidates move along the left-right spectrum, enabling us to determine whether a 20-point change in CPI corresponds to meaningful differences in how candidates discuss abortion.

These analyses reveal consistent patterns that confirm the substantive magnitude of average positional shifts across issues. Table 3 displays the top fifteen discriminating words identified through keyness analyses for each percentile comparison; results for all other CPI issue areas are included in Appendix Tables A24 to A28. Candidates in higher-percentile groups more frequently use terms such as *life, unborn, conception,* and *pro-life*. Conversely, lower-percentile candidates more often

[19]Each percentile group includes issue statements within a 10-point range of that cutoff (e.g., the "90th" percentile group includes candidates ranked from the 85th to 95th percentile). This ensures sufficient sample sizes while maintaining an average 20-point gap between groups.

Table 3: Abortion Word Exclusivity Comparison: CPI-Abortion

Percentile Comparison	**Higher Percentile** *Right-Leaning*	**Lower Percentile** *Left-Leaning*
90th vs. 70th	life, conception, begins, every, gift, protect, natural, fight, lives, innocent, human, precious, god, sanctity, believe	abortion, health, healthcare, care, decision, also, roe, wade, legislation, pregnancy, v, people, federal, supported, family
70th vs. 50th	life, unborn, conception, begins, protect, taxpayer, innocent, congress, congressman, proud, god, vote, defund, bill, pro-life	women, access, control, decision, health, reproductive, woman, decisions, personal, trimester, court, state, doctor, options
50th vs. 30th	life, pro-life, unborn, mother, abortions, conception, abortion, children, pregnancy, may, rape, adoption, baby, believe, trimester	reproductive, access, health, care, women's, services, equal, congress, women's, rights, women, safe, woman's, healthcare, planning
30th vs. 10th	bill, life, abortions, voted, republican, assistance, continue, abortion, want, working, court, cancer, stand, nation, congress	hyde, amendment, ensure, reproductive, healthcare, comprehensive, repeal, contraception, workplace, trans, women, status

employ language centered on *healthcare, access,* and *rights*. Importantly, if 20-percentile differences were merely measurement error, we would expect to see random or nonspecific word use between percentile groups rather than these systematically meaningful political distinctions. The recurring appearance of politically coherent terminology suggests that these positional shifts reflect genuine substantive differences in candidate left-right positioning, not just statistical noise.

District Policy Opinion & Candidate Policy-Level Positioning

Having established that candidates adopt meaningfully different left-right stances across issue areas, we analyze whether these positions correspond with district-level policy preferences. Previous research shows that voters and constituencies hold inconsistent left-right preferences across issues (Warshaw and Rodden 2012; Broockman 2016). Under the assumption that candidates adopt left-right positions that maximize their electoral appeal, we argue that candidates' issue-specific positions should reflect the diversity of their district's issue-specific preferences. We test this expectation with a series of linear regressions.

Our main explanatory variables in all models are district-level public opinion measures on specific policy proposals. We primarily rely on estimates from de Benedictis-Kessner et al. (2024), who use policy preference questions from 18 large-scale surveys of the American public to esti-

Table 4: District-Level Policy-Specific Survey Questions

Issue Area	**Survey Question: Do you support or oppose...?**
Abortion	*Pro-Choice* (L): Always allowing a woman to obtain an abortion as a matter of choice *Pro-Life* (R): Make abortions illegal in all circumstances
Education	*Local* (R): Provide tax-funded vouchers to be used for private or religious schools
Energy	*Renewables* (L): Fund research into renewable energy sources *Fossil Fuels* (R): Expand offshore drilling for oil and natural gas off the U.S. coast
Guns	*Restrictions* (L): Banning assault rifles *Restrictions* (L): Require criminal background checks on all gun sales
Healthcare	*Publicly Funded* (L): Providing Medicare (Medicare-for-all) for all Americans *Privatized* (R): Repealing the entire Affordable Care Act
Immigration	*Inclusive* (L): Provide permanent resident status to children of immigrants *Exclusive* (R): Deport all undocumented immigrants

mate public opinion on issues at the congressional district level.[20] District-level estimates for the 2018 and 2020 election cycles are based on surveys fielded between 2017 and 2020; estimates for the 2022 cycle are based on surveys fielded between 2021 and 2023. To generate their estimates, de Benedictis-Kessner et al. (2024) use an estimation strategy called multilevel regression and post-stratification (MRP), which models individual survey responses as a function of demographic and geographic variables, then adjusts these estimates using population data to produce predictions. We supplement these data with district-level estimates of energy policy preferences from Howe et al. (2015), who use a comparable MRP approach. Table 4 shows the wording of the policy questions used in our analyses. We combine the public opinion estimates derived from these questions by policy area to create a single summary measure for each issue-district-year combination.[21]

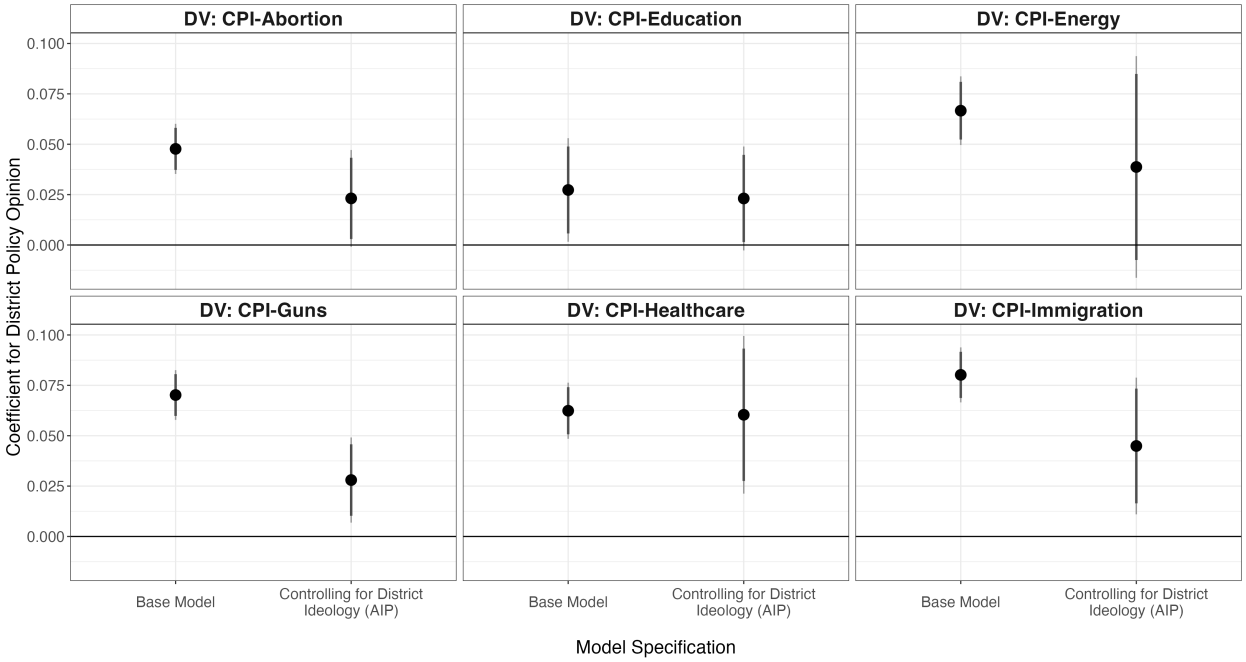
For each issue area, we estimate two linear models. In the first, the dependent variable is a candidate's issue-specific CPI.[22] Model predictors include the district-level public opinion measure for that issue area and a set of candidate-level and district-level controls (e.g., partisanship, previous electoral experience, and primary type). If candidates' policy positioning reflect district

[20]Specific data sources include the 2009-2023 Cooperative Election Studies (Vavreck and Rivers 2008; Ansolabehere and Rivers 2013) and the 2019-2021 UCLA/Nationscape Surveys (Tausanovitch et al. 2019).

[21]To create these measures, we first reorient the component estimates to ensure they point in the same direction. We then rescale each to have a mean of 0 and a standard deviation of 1, and combine them by addition—giving equal weight to each component. Our main results are robust to using disaggregated estimates.

[22]We exclude candidates who did not discuss a given issue and therefore have no issue-specific CPI generated.

Figure 7: OLS Regression Model Coefficients: District Policy-Level Opinion and CPIs



Note: Panels display model coefficients. The left coefficients are drawn from base models (Appendix Table A29). The right coefficients are drawn from models replicating these analyses, controlling for district overall ideological orientation (Appendix Table A30). Bars indicate 95% and 90% confidence intervals.

policy preferences, we should observe a statistically significant, positive relationship between public opinion estimates and issue-related CPIs. In the second model, we introduce an additional control for a district's overall ideological orientation, using ideal point estimates from Warshaw and Tausanovitch (2022). These estimates are constructed via a two-stage procedure: individual-level ideology is first estimated using a Bayesian item-response theory (IRT) model, and then aggregated to the district level using MRP. Unlike the policy-specific opinion measures described above, these ideal points summarize each district's overall ideology across a broad set of issues. This allows us to evaluate whether candidates' policy-level positions align with issue-specific public opinion beyond what would be expected based on the overall ideological leaning of their district.

Figure 7 presents estimated coefficients from linear regression models, with issue-specific CPI as the outcome of interest. The left point estimates in each panel denote model coefficients for district policy opinion in baseline models, which include the candidate and district controls discussed above. The right point estimates denote coefficients in models that additionally control for

the overall ideological orientation of districts. Point estimates show coefficient values, with error bars representing 95% and 90% confidence intervals. Full model outputs are provided in Appendix Tables A29 and A30.

Results in Figure 7 show a clear link between the average district-specific policy opinions of constituents and issue-specific positioning by U.S. House candidates. In all six baseline models (Figure 7, left coefficients), we find positive and statistically significant relationships between district-level policy opinion and candidates' issue-related CPIs at the .05 level, indicating that candidate policy positions generally reflect local constituency preferences. This relationship persists even after accounting for overall district ideology (Figure 7, right coefficients), with district policy opinion continuing to be a significant predictor of issue-related CPI in three models at the .05 level (guns, healthcare, and immigration) and in two models at the .10 level (abortion and education). Only the model with CPI-Energy as the outcome falls short of traditional significance thresholds ($p = 0.17$). Importantly, district policy-specific opinions are not merely substituting for overall district ideology as a predictor of candidate positioning; these ideal point measures of district political orientation are statistically significant at the 0.05 level in four of the six models, underscoring the independent contributions of both estimates for districts' ideological leanings.

The persistence of issue-specific public opinion effects in our results, beyond overall district ideology, suggests that candidates' left-right stances on policy track with district issue preferences on these same issues. When we replicate these analyses using unidimensional measures of overall candidate positioning from Bonica (2024) and Case (2025), we do not observe the same consistent, statistically significant relationships as in Figure 7.[23] In some cases, we observe notable *negative* relationships, meaning that as districts become more left-leaning on a specific policy area, candidates tend to adopt more right-leaning overall ideologies. This pattern suggests that unidimensional scores can mask important variations in how candidates position themselves on issues relative to district opinion, leading to different conclusions about democratic representation.

[23]These results are shown in Appendix Figure A4.

Ideological Constraint & Electoral Competition

Our final empirical analyses investigate whether candidates exhibit greater cross-issue heterogeneity in their policy positions when running in electorally unfavorable districts. We theorize that positional inconsistency should be most evident in districts where candidates cannot rely exclusively on co-partisan support and must appeal to a broader coalition, including cross-pressured or swing voters. In such contexts, adopting divergent left-right stances across different issue domains may serve as a strategic tool to broaden electoral appeal. By contrast, in safe districts where candidates can rely on their partisan base, the incentive to adopt differentiated positions across issues is substantially reduced. Candidates may be reticent to take diverging extreme and moderate stances in same-party safe seats, fearing that they will alienate politically knowledgeable and ideologically extreme primary voters.

Although previous research has examined the electoral effects of positional inconsistency (Rogowski and Tucker 2018; Somer-Topcu 2015), these studies primarily rely on indirect proxies rather than direct measures of cross-issue ideological variation. Our analysis advances this line of work by directly measuring candidates' stated policy positions across multiple issue areas, enabling a more precise assessment of when and why candidates diverge from ideologically consistent stances. Importantly, our analytical focus centers on the strategic determinants of ideological diversity rather than its electoral consequences, providing new insights into how electoral contexts influence candidate behavior.

To examine the relationship between candidate-level positional constraint and district-level partisan competition, we estimate a series of ordinary least squares (OLS) models. The dependent variable in all model specifications is within-candidate variance in CPI scores—a measure adapted from Rogowski and Tucker (2018) to measure cross-issue differences in candidates' stated positions.[24] Our main explanatory variables include measures for the electoral security of a district for the candidate's party. We operationalize district-level partisan competition using two indicators:

[24]We limit our sample to office-seeking candidates with CPI scores on at least two issue areas within a given election year, as at least two observations are needed to calculate within-candidate variance in policy positions.

Table 5: Model Results: Ideological Constraint and Partisan Competitiveness

	DV: Variance in CPIs			
	(1)	(2)	(3)	(4)
Same Party Vote	-0.0002^{***}	-0.0003^{*}		
	(0.00004)	(0.0002)		
Same Party Ideology			-0.017^{***}	-0.031^{*}
			(0.003)	(0.017)
Partisan Primary	-0.003^{*}		-0.002^{*}	
	(0.002)		(0.002)	
Fundraising (Logged)	-0.0004^{***}		-0.0004^{***}	
	(0.0001)		(0.0001)	
Observations	3,965	3,965	3,965	3,965
Candidate Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Note:				*p<0.1; **p<0.05; ***p<0.01

(1) the two-party vote share received by the candidate's co-partisan in the most recent presidential election, and (2) district-level overall ideological orientation estimates from Warshaw and Tausanovitch (2022).

For each explanatory variable, we estimate both cross-sectional and two-way fixed effects (TWFE) models. Cross-sectional regressions capture between-candidate variation in a given election cycle, offering a broad view of how district partisan dynamics correlate with positional heterogeneity, as well as relationships with other candidate and district-level factors. In particular, we control for candidate fundraising, partisanship, gender, and prior elected experience, as well as district primary election rules, election year, and open-seat status. TWFE models, by contrast, leverage within-candidate changes over time, controlling for unobserved candidate-specific traits and year-specific shocks.

Table 5 presents truncated results from four regression models examining the relationship between electoral context and candidates' ideological consistency. Full results are available in Ap-

pendix Table A31. Recall that we measure ideological consistency using the variance in candidates' CPIs, where lower variance indicates greater constraint or consistency in left-right positioning across issues. We find that candidates in safer seats are more ideologically constrained. In Models 1 and 2, the coefficient for *Same Party Vote* is negative and statistically significant, meaning that as the candidate's party's presidential vote share in a district increases, candidates tend to adopt more consistent left-right positions across issues. This suggests that electoral security decreases incentives for cross-issue differentiation, aligning with our expectations. Models 3 and 4 use our alternative measure of district-level partisan competition. The negative and statistically significant coefficient for *Same Party Ideology* indicates that Democratic (Republican) candidates running in districts that are more left-leaning (right-leaning) maintain more constrained ideological positioning. Together, these findings support the conclusion that partisan electoral competition influences the degree of ideological consistency among candidates across the policy positions they adopt in their campaigns.

Two additional patterns emerge from our analysis that warrant discussion. Across cross-sectional models, candidates in partisan primaries also exhibit significantly lower variance in CPI scores, suggesting that intraparty competition is associated with greater consistency in issue positioning. This result challenges the finding that primary election institutions have little influence on elite positioning (see McGhee et al. 2014). Finally, higher logged pre-primary fundraising totals are consistently associated with greater ideological constraint. Well-funded candidates may be better resourced to craft cohesive messaging strategies or may attract donors who reward ideological consistency (Porter et al. 2024; Porter and Treul 2025)

Discussion & Conclusion

This paper questions a fundamental assumption in the study of American politics: that congressional candidates hold consistent left-right positions across issue areas. To advance our understanding of candidate positioning, this article introduces a novel text-to-measure pipeline for extracting and scaling latent policy positions within texts. We additionally employ an original compilation of campaign platforms sourced from congressional candidates' campaign websites to create our

positioning estimates. The text data introduced in this paper alone represents a significant contribution and will be made open-source to advance the study of representation in American legislative politics. We produce estimates of policy-level positioning, which we refer to as Candidate Positioning Indexes (CPIs), for three-quarters of all major-party, ballot-eligible candidates who ran in primaries for the U.S. House between 2018 and 2022. Our measurement focuses on major policy cleavages across six policy areas that voters expressed as pivotal to guiding their vote choice in recent elections: abortion, education, energy, guns, healthcare, and immigration. Importantly, our measurement strategy is flexible and can be adapted for various other policy applications within US politics and beyond.

Three key findings emerge from our analysis. First, the modest pairwise correlations between Candidate Positioning Estimates across issues suggest that ideological consistency is not as common as existing expectations would indicate. On average, candidates shift 20 percentile ranks in relative left-right positioning from one issue to another—a level of variation that we show to be substantively meaningful through both qualitative and quantitative assessments. Second, this variation in positioning aligns with district-level, issue-specific public opinion, suggesting that candidates strategically tailor their stances to local preferences on individual policies. This finding challenges traditional models of representation that assume candidates maintain uniform ideological positions and instead supports the idea of issue-specific district representation. Third, and perhaps most importantly, positional inconsistency increases with district partisan insecurity: candidates in districts that do not favor their party adopt more varied positions, potentially to broaden their electoral appeal to cross-pressured and swing voters.

These findings appear to conflict with a large amount of research showing ideological constraint in legislators' roll-call behavior. One explanation for this divergence in findings lies in the distinct institutional settings of campaigning versus governing. Roll-call votes are shaped by strong party leadership, agenda control, and coalition maintenance, all of which exert pressure toward party-line voting and ideological conformity (Pearson 2015; Snyder and Groseclose 2000). By contrast, campaign platforms give candidates considerable autonomy to focus on issue positions that appeal

to local constituencies, without the formal constraints of party discipline or legislative bargaining. Work by Marble and Tyler (2022) captures this disconnect, finding much less ideological constraint and greater multidimensionality in positioning estimates based on roll-call data compared to surveys of candidates.

Our findings also speak to ongoing debates about the nature of political polarization in America. Rather than viewing polarization as a uniform phenomenon across all policy domains, our results suggest that the intensity of elite issue polarization varies considerably by policy area and electoral context. Competitive districts, in particular, appear to incentivize a form of strategic positioning that operates at the issue level rather than through blanket ideological centrism. This insight offers a more optimistic view of democratic responsiveness: even in a polarized era, electoral competition can promote substantive representation on specific policies that matter to local constituencies.

Several limitations of our approach point toward productive avenues for future research. While we focus on six major issue areas, expanding this analysis to additional policy domains could reveal further patterns in strategic positioning. Moreover, our cross-sectional design limits our ability to causally trace how individual candidates adjust their positions over time or in response to changing electoral conditions. Future work might also explore how voters perceive and respond to positional inconsistency, and whether the electoral benefits of strategic positioning outweigh potential costs in terms of perceived authenticity or ideological clarity.

This study demonstrates that candidate positioning involves a complex calculation of partisan expectations, local preferences, and competitive pressures rather than solely ideological or opportunistic motives. By clarifying the conditions under which candidates deviate from ideological consistency, we highlight ways in which democratic institutions can encourage meaningful policy representation, even in an increasingly nationalized political environment. As American democracy grapples with deepening partisan divides, understanding these mechanisms of strategic moderation becomes increasingly important for finding opportunities for policy compromise and democratic renewal.

References

Abbe, Owen G., Jay Goodliffe, Paul S. Herrnson, Kelly D. Patterson, and Kelly D. Patterson. 2003. "Agenda Setting in Congressional Elections: The Impact of Issues and Campaigns on Voting Behavior." *Political Research Quarterly* 56(4): 419–430.

Ahler, Douglas J., and David E. Broockman. 2018. "The Delegate Paradox: Why Polarized Politicians Can Represent Citizens Best." *Journal of Politics* 80(4).

Aldrich, John H., Jacob M. Montgomery, and David B. Sparks. 2014. "Polarization and ideology: Partisan sources of low dimensionality in scaled roll call analyse." *Politial Analysis* 22(4): 435–456.

Alvarez, Michael R. 1997. *Information & Elections*. University of Michigan Press.

Anderson, Sarah E., Daniel M. Butler, and Laurel Harbridge-Yong. 2024. "The Primary Premium: Why Legislators are Incentivized to Side with Primary Voters over General Election Voters.".

Ansolabehere, Stephen, and Douglas Rivers. 2013. "Cooperative Survey Research." *Annual Review of Political Science* 16(1): 307–329.

Ansolabehere, Stephen, James M. Snyder Jr., and Charles Stewart III. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45(1).

Bafumi, Joseph, and Michael C. Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress." *American Political Science Review* 104(3): 519–542.

Bailey, Michael A. 2024. "Measuring Candidate Ideology in Congressional Elections, 2020-2022.".

Baker, Anne E. 2016. "Do Interest Group Endorsements Cue Individual Contributions to House Candidates?" *American Politics Research* 44(2).

Barber, Michael. 2016. "Donation Motivations: Testing Theories of Access and Ideology." *Political Research Quarterly* 69(1): 148–159.

Barber, Michael, and Jeremy C Pope. 2018. Who is ideological? measuring ideological consistency in the american public. In *The Forum*. Vol. 16 De Gruyter pp. 97–122.

Barbera, Pablo. 2015. "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data." *American Political Science Review* 23(1): 76–91.

Bateman, David A., Joshua D. Clinton, and John S. Lapinski. 2017. "A House Divided? Roll Calls, Polarization, and Policy Differences in the U.S. House, 1877–2011." *American Journal of Political Science* 61(3): 698–714.

Bolukbasi, Tolga, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. "Man is to computer programmer as woman is to homemaker? debiasing word embeddings." *Advances in neural information processing systems* 29.

Bonica, Adam. 2013. "Ideology and interests in the political marketplace." *American Journal of Political Science* 2(57): 294–311.

Bonica, Adam. 2014. "Mapping the Ideological Marketplace." *American Journal of Political Science* 58(2): 367–386.

Bonica, Adam. 2024. "Database on Ideology, Money in Politics, and Elections: Public version 4.0 [Computer file].".

Bonica, Adam, and Zhao Li. 2021. "Campaign Agendas and Issue Group Strategy in Congressional Primaries.".

Brady, David W., Hahrie Han, and Jeremy C. Pope. 2007. "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" *Legislative Studies Quarterly* 32(1): 79–105.

Broockman, David E. 2016. "Approaches to studying policy representation." *Legislative Studies Quarterly* 41(1): 181–215.

Broockman, David E., and Benjamin E. Lauderdale. 2025. ""Moderates"." *American Political Science Review* .

Burden, Barry C. 2004. "Candidate Positioning in US Congressional Elections." *British Journal of Political Science* 34: 211–227.

Carson, Jamie L., Jason M. Roberts, and Rachel Porter. 2020. "Congress and the Nationalization of Congressional Elections." In *New Directions in Congressional Politics*. Routledge.

Case, Colin R. 2025. "Measuring Strategic Positioning in Congressional Elections." *Journal of Politics* .

Clinton, Joshua D. 2012. "Using roll call estimates to test models of politics." *Annual Review of Political Science* 15(1): 79–99.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll-Call Data." *American Political Science Review* 98(May): 355–370.

Colao, Jacqueline, David E. Broockman, Gregory A. Huber, and Joshua L. Kalla. 2025. "Tracing Polarization's Roots: A Panel Study of Voter Choice in Congressional Primary and General Elections.".

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." *Critical Review* (1-74).

Crespin, Michael H., and David W. Rohde. 2010. "Dimensions, issues, and bills: Appropriations voting on the House floor." *Journal of Politics* 72(4): 976–989.

Curry, James M., and Frances E. Lee. 2020. *The Limits of Party: Congress and Lawmaking in a Polarized Era*. University of Chicago Press.

Daubler, Thomas, Kenneth Benoit, Slava Mikhaylov, and Michael Laver. 2012. "Natural sentences as valid units for coded political texts." *British Journal of Political Science* 42(4): 937–951.

de Benedictis-Kessner, Justin, Ruth Greenwood, Nicholas Stephanopoulos, and Christopher Warshaw. 2024. "TrueViews–See What People Think.".

Denny, Matthew J., and Arthur Spirling. 2018. "Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It." *Political Analysis* 26(2): 168–189.

Dolan, Kathleen. 2005. "Do women candidates play to gender stereotypes? Do men candidates play to women? Candidate sex and issues priorities on campaign websites." *Political Research Quarterly* 58(1): 31–44.

Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy* 65: 135–150.

Druckman, James N., Martin J. Kifer, and Michael Parkin. 2009. "Campaign Communications in U.S. Congressional Elections." *American Political Science Review* 103(3): 343–366.

Egan, Patrick J. 2013. *Partisan priorities: How issue ownership drives and distorts American politics*. Cambridge University Press.

Egerod, Benjamin CK, and Robert Klemmensen. 2020. "Scaling political positions from text: Assumptions, Methods and Pitfalls." In *he SAGE Handbook of Research Methods in Political Science and International Relations*. SAGE Publications pp. 498–521.

Fenno, Richard F. 1978. *Home Style: House Members in Their Districts*. New York: Longman.

Fouirnaies, Alexander, and Andrew B. Hall. 2014. "The Financial Incumbency Advantage: Causes and Consequences." *Journal of Politics* 76(3): 711–724.

Fowler, Anthony, and Andrew Hall. 2016. "The Elusive Quest for Convergence." *Quarterly Journal of Political Science* 11(1): 131–149.

Freeder, Sean, Gabriel S. Lenz, and Shad Turney. 2022. "The importance of knowing "what goes with what": Reinterpreting the evidence on policy attitude stability." *Journal of Politics* 81(1): 274–290.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. "Word embeddings quantify 100 years of gender and ethnic stereotypes." *Proceedings of the National Academy of Sciences* 115(16): E3635–E3644.

Garten, Justin, Joe Hoover, Kate M Johnson, Reihane Boghrati, Carol Iskiwitch, and Morteza Dehghani. 2018. "Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis: Distributed dictionary representation." *Behavior research methods* 50: 344–361.

Grand, Gabriel, Idan Asher Blank, Francisco Pereira, and Evelina Fedorenko. 2022. "Semantic projection recovers rich human knowledge of multiple object features from word embeddings." *Nature Human Behavior* 6: 975–987.

Grimmer, Justin, and Brandon M Stewart. 2013. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political analysis* 21(3): 267–297.

Grimmer, Justin, Margaret E. Roberts, and Brandon M. Stewart. 2022. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press.

Hall, Andrew B. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review* 109(1): 18–42.

Harden, Jeffrey J. 2015. *Multidimensional democracy: A supply and demand theory of representation in American legislatures*. Cambridge University Press.

Hayes, Danny, and Jennifer L. Lawless. 2016. *Women on the Run: Gender, Media, and Political Campaigns in a Polarized Era*. Cambridge University Press.

Henderson, John A., Geoffrey Sheagley, Stephen N. Goggin, Logan Dancey, and Alexander G. Theodoridis. 2022. "Primary divisions: How voters evaluate policy and group differences in intraparty contests." *Journal of Politics* 84(3): 1760–1776.

Hill, Seth J. 2015. "Institution of nomination and the policy ideology of primary electorates." *Quarterly Journal of Political Science* 10(4).

Hopkins, Daniel J. 2018. *The Increasingly United States: How and Why American Political Behavior Nationalized*. University of Chicago Press.

Howe, Peter D., Matto Mildenberger, Jennifer R. Marlon, and Anthony Leiserowitz. 2015. "Geographic Variation in Opinions on Climate Change at State and Local Scales in the USA." *Nature Climate Change* 5(6): 596–603.

Jacobson, Gary C., and Jamie L. Carson. 2016. *The Politics of Congressional Elections*. 7th edition ed. New York: Longman.

Jochim, Ashley E., and Bryan D. Jones. 2013. "Issue politics in a polarized congress." *Political Research Quarterly* 66(2): 352–369.

Kaplan, Noah, David K. Park, and Travis N. Ridout. 2006. "Dialogue in American Political Campaigns? An Examination of Issue Convergence in Candidate Television Advertising." *American Journal of Political Science* 50(3): 724–736.

Kingdon, John W. 1973. *Congressman's Voting Decisions*. Harper and Row.

Kitagawa, Risa, and Fiona Shen-Bayh. 2024. "Measuring Political Narratives in African News Media:A Word Embeddings Approach." *Journal of Politics* .

Kozlowski, Austin C., Matt Taddy, and James A. Evans. 2019. "The Geometry of Culture: Analyzing Meaning through Word Embeddings." *American Sociological Review* 84(5): 905–949.

Lapinski, John, Matt Levendusky, Ken Winneg, and Kathleen Hall Jamieson. 2016. "What do citizens want from their member of congress?" *Political Research Quarterly* 69(3): 535–545.

Lau, Jey Han, and Timothy Baldwin. 2016. "An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation." *1st Workshop on Representation Learning for NLP* .

Lauderdale, Benjamin E. 2010. "Unpredictable voters in ideal point estimation." *Political Analysis* 18(2): 151–171.

Lauderdale, Benjamin E., Chris Hanretty, and Nick Vivyan. 2018. "Decomposing Public Opinion Variation into Ideology, Idiosyncrasy, and Instability." *Journal of Politics* 80(2): 707–712.

Le, Quoc, and Tomas Mikolov. 2014. Distributed representations of sentences and documents. In *International conference on machine learning*. PMLR pp. 1188–1196.

Lee, Frances E. 2016. *Insecure Majorities: Congress and the Perpetual Campaign*. Chicago, IL: University of Chicago Press.

Lee, Francs E. 2015. "How party polarization affects governance." *Annual Review of Political Science* 18(1): 261–282.

Levendusky, Matthew. 2023. *Our Common Bonds: Using What Americans Share to Help Bridge the Partisan Divide*. Chicago University Press.

Li, Zhao. 2018. "How Internal Constraints Shape Interest Group Activities: Evidence from Access-Seeking PACs." *American Political Science Review* 112(4): 792–808.

Lowe, Will, and Kenneth Benoit. 2013. "Validating Estimates of Latent Traits from Textual Data Using Human Judgment as a Benchmark." *British Journal of Political Science* 21: 298–313.

Macdonald, Maggie, Megan A. Brown, Joshua A. Tucker, and Jonathan Nagler. 2025. "To moderate, or not to moderate: Strategic domain sharing by congressional campaigns." *Electoral Studies* 95.

Marble, William, and Matthew Tyler. 2022. "The Structure of Political Choices: Distinguishing Between Constraint and Multidimensionality." *American Journal of Political Science* 30: 328–345.

Mayhew, David R. 1974. *Congress: The Electoral Connection*. New Haven: Yale University Press.

McGhee, Eric, Seth Masket, Boris Shor, Steven Rogers, and Nolan McCarty. 2014. "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology." *American Journal of Political Science* 58(2).

Meisels, Mellissa. 2024. "Campaign Agendas and Issue Group Strategy in Congressional Primaries.".

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient estimation of word representations in vector space." *arXiv preprint arXiv:1301.3781* .

Montagnes, Brendan Pablo, and Jon C. Rogowski. 2015. "Testing Core Predictions of Spatial Models: Platform Moderation and Challenger Success." *Political Science Research and Methods* 3(3): 619–640.

Moser, Scott, Abel Rodríguez, and Chelsea L. Lofland. 2021. "Multiple ideal points: Revealed preferences in different domains." *American Journal of Political Science* 29(2): 139–166.

Moskowitz, Daniel J., Jon C. Rogowski, and James M. Snyder Jr. 2024. "Parsing Party Polarization in Congress." *Quarterly Journal of Political Science* .

Park, Ju Yeon, and Jacob M. Montgomery. 2023. "Towards a framework for creating trustworthy measures with supervised machine learning.".

Pearson, Kathryn. 2015. *Party Discipline in the U.S. House of Representatives*. University of Michigan Press.

Pennec, Caroline Le, Rafael Di Tella, Randy Kotti, and Vincent Pons. 2024. "Keep your Enemies Closer: Strategic Candidate Adjustments in U.S. and French Elections.".

Petrocik, John R., William L. Benoit, and Glenn J. Hansen. 2003. "Issue ownership and presidential campaigning, 1952–2000." *Political Science Quarterly* 118(4): 599–626.

Poole, Keith T., and Howard L. Rosenthal. 2011. *Ideology and congress*. Transaction Publishers.

Poole, Keith T., and Howard Rosenthal. 2000. *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press.

Poole, Keith T., and R. Steven Daniels. 1985. "Ideology, Party, and Voting in the U.S. Congress, 1959–1980." *American Political Science Review* 72(2): 373–399.

Poole, Keith T., Thomas Romer, and Howard Rosenthal. 1987. "The Revealed Preferences of Political Action Committees." *The American Economic Review* 77(2): 298–302.

Porter, Rachel. 2022. Some Politics Are Still Local: Strategic Position Taking in Congress and Elections Phd thesis, University of North Carolina at Chapel Hill.

Porter, Rachel, and Sarah A. Treul. 2025. "Evaluating (In)Experience in Congressional Elections." *American Journal of Political Science* 69(1): 284–298.

Porter, Rachel, Sarah A. Treul, and Maura McDonald. 2024. "Changing the Dialogue: Descriptive Candidacies and Position Taking in Campaigns for the US House of Representatives." *Journal of Politics* .

Porter, Rachel, William Kakenmaster, and Ben Francis. 2025. "Roads to Climate Compromise in the US Congress and Congressional Elections.".

Powell, Lynda W. 1982. "Issue Representation in Congress." *The Journal of Politics* 44(3): 658–678.

Rheault, Ludovic, and Christopher Cochrane. 2020. "Word embeddings for the analysis of ideological placement in parliamentary corpora." *Political Analysis* 28(1): 112–133.

Roberts, Jason M., and Steven S. Smith. 2003. "Procedural Contexts, Party Strategy, and Conditional Party Voting in the U.S. House of Representatives, 1971-2000." *American Journal of Political Science* 47(April): 305–317.

Roberts, Jason M., Steven S. Smith, and Stephen R. Haptonstahl. 2016. "The Dimensionality of Congressional Voting Reconsidered." *American Politics Research* 44(5): 794–815.

Rodriguez, Pedro L., and Arthur Spirling. 2022. "Word embeddings: What works, what doesn't, and how to tell the difference for applied research." *Journal of Politics* 84(1): 101–115.

Rogowski, Jon C, and Patrick D Tucker. 2018. "Moderate, extreme, or both? How voters respond to ideologically unpredictable candidates." *Electoral Studies* 51: 83–92.

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

Ryan, Timothy J., and J. Andrew Ehlinger. 2023. *Issue Publics: How Electoral Constituencies Hide in Plain Sight*. Cambridge University Press.

Shin, Sooahn. 2024. "Measuring Issue Specific Ideal Points from Roll Call Votes.".

Snyder, James M., and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science* 46(1): 90–110.

Snyder, James M., and Timothy Groseclose. 2000. "Estimating Party Influence in Congressional Roll-Call Voting." *American Journal of Political Science* 44(2): 193–211.

Somer-Topcu, Zeynep. 2015. "Everything to Everyone: The Electoral Consequences of the Broad-Appeal Strategy in Europe." *American Political Science Review* 59(4): 841–854.

Stone, Walter J., and Elizabeth N. Simas. 2010. "Candidate Valence and Ideological Positions in U.S. House Elections." *American Journal of Political Science* 54(2): 371–388.

Sulkin, Tracy. 2005. *Issue Politics in Congress*. Cambridge University Press.

Sulkin, Tracy. 2011. *The Legislative Legacy of Congressional Campaigns*. Cambridge University Press.

Sulkin, Tracy, Cortney M. Moriarty, and Veronica Hefner. 2007. "Congressional Candidates' Issue Agendas On- and Off-line." *The International Journal of Press/Politics* 12(2): 63–79.

Tausanovitch, Chris, and Christopher Warshaw. 2017. "Estimating Candidates' Political Orientation in a Polarized Congress." *Political Analysis* 25(2): 167–187.

Tausanovitch, Chris, Lynn Vavreck, Tyler Reny, Alex Rossell Hayes, and Aaron Rudkin. 2019. Democracy Fund+ UCLA Nationscape Methodology and Representativeness Assessment. Report Democracy Fund Voter Study Group.

Tomz, Michael, and Robert P. Van Houweling. 2009. "The electoral implications of candidate ambiguity." *American Political Science Review* 103: 83–98.

Treier, Shawn, and D. Sunshine Hillygus. 2009. "The nature of political ideology in the contemporary electorate." *Public Opinion Quarterly* 73(4): 679–703.

Vavreck, Lynn, and Douglas Rivers. 2008. "The 2006 Cooperative Congressional Election Study." *Journal of Elections, Public Opinion and Parties* 18(4): 355–366.

Warshaw, Christopher, and Chris Tausanovitch. 2022. "Subnational Ideology and Presidential Vote Estimates (v2022)." https://doi.org/10.7910/DVN/BQKU4M.

Warshaw, Christopher, and Jontahan Rodden. 2012. "How Should We Measure District-Level Public Opinion on Individual Issues." *Journal of Politics* 74(1): 203–19.

Xenos, Michael A., and Kirsten A. Foot. 2005. "Politics As Usual, or Politics Unusual? Position Taking and Dialogue on Campaign Websites in the 2002 U.S. Elections." *Journal of Communication* 55(1): Pages 169–185.

Yoshinaka, Antoine, and Christian R. Grose. 2011. "Ideological hedging in uncertain times: Inconsistent legislative representation and voter enfranchisement." *British Journal of Political Science* 41(4): 765–794.

You, Hye Young. 2023. "Dynamic lobbying: Evidence from foreign lobbying in the US Congress." *Economics & Politics* 35(2): 445–469.

Strategic Heterogeneity in Policy-Level Positioning: Evidence from Congressional Campaigns (Supplementary Materials)

July 8, 2025

Contents

A1 Data Collection

To collect text data from candidate campaign websites, we first identify the names of all major party candidates running in 2018, 2020, and 2022 using candidate filings with the Federal Election Commission (FEC) and state-level elections websites. Using this list of names, we seek to identify the campaign website URLs for all candidates in each election year by following links from online repositories like Politics1.com, visiting candidates' social media pages, and conducting simple Google searches. Candidates for whom we locate no campaign website are re-checked again in the days leading up to their primary to ensure we did not miss their site. Using this approach, we identify 5,478 of 6,006 major-party, ballot-eligible candidates who ran for the US House of Representatives between 2018 and 2022 as having a campaign website.

Once we identify the URL for a candidate's campaign website, we determine whether that candidate also had a policy platform on her website. We trace the process of collecting text from campaign platforms in Figure A1. For many candidates, this was a simple process; platform pages often had clear titles like "Where I Stand" or "My Positions." Of those 5,478 congressional candidates who had a campaign website, 82% also had a campaign platform—or 75% of all major-party, ballot-eligible candidates who ran between 2018 and 2022. In Table A1, we explore predictors for missingness in our data. We estimate a linear probability model (left column) and logistic regression (right column) where the outcome variable is whether (1) or not (0) a given candidate had a campaign website with a policy platform. We find that running unopposed in a primary, being a Republican candidate, running in 2020, and being an office-holder are all statistically significant predictors for not having a campaign platform.

After locating a campaign website (Figure A1, panel 1) and campaign platform (Figure A1, panel 2), we locate the text for platform points. In nearly all cases, text presented on a candidate's campaign platform is organized as a series of platform points. We define a platform point as the body text nested under a descriptive text header. For some candidates, platform point text is included on a single page as a series of paragraphs broken up by headers denoting the issue area for that position (e.g., Abortion, Second Amendment Rights, or Agriculture). In the case shown in panel 3 of Figure A1, each platform point has its own dedicated sub-page.

We archive text from campaign platforms using a Qualtrics form. First, we record the unique identifier for the candidate (Figure A1, panel 4). Next, we collect meta-data for the candidate (Figure A1, panel 5), such as their race, gender, past elected experience, and fundraising identifier from the Federal Election Commission (FEC). We collect this information from a candidate's website; we rely on auxiliary resources (e.g., state legislative biographic profiles, newspaper articles, Wikipedia summaries, and Ballotpedia profiles) when needed. Finally, we archive text for each platform point (Figure A1, panel 6).

To ensure consistency in text collection, we scrape campaign websites during a two-week window before a candidate's primary election. We collect text before the primary because many candidates deactivate their websites immediately after losing their race. Internet archive repositories, such as the Way Back Machine or Library of Congress, catalog the campaign websites for some congressional candidates. However, many candidates' websites, particularly primary candidates' websites, are not archived. Further, text from "Issues" pages on archived sites is often not cataloged consistently. Web-scraping crawlers for the Way Back Machine only catalog for specific URLs at a fixed time. For example, if "https://www.ocasiocortez.com/" is scraped on a given day, this does not mean sub-pages on this site, such as "https://www.ocasiocortez.com/issues," are also scraped.

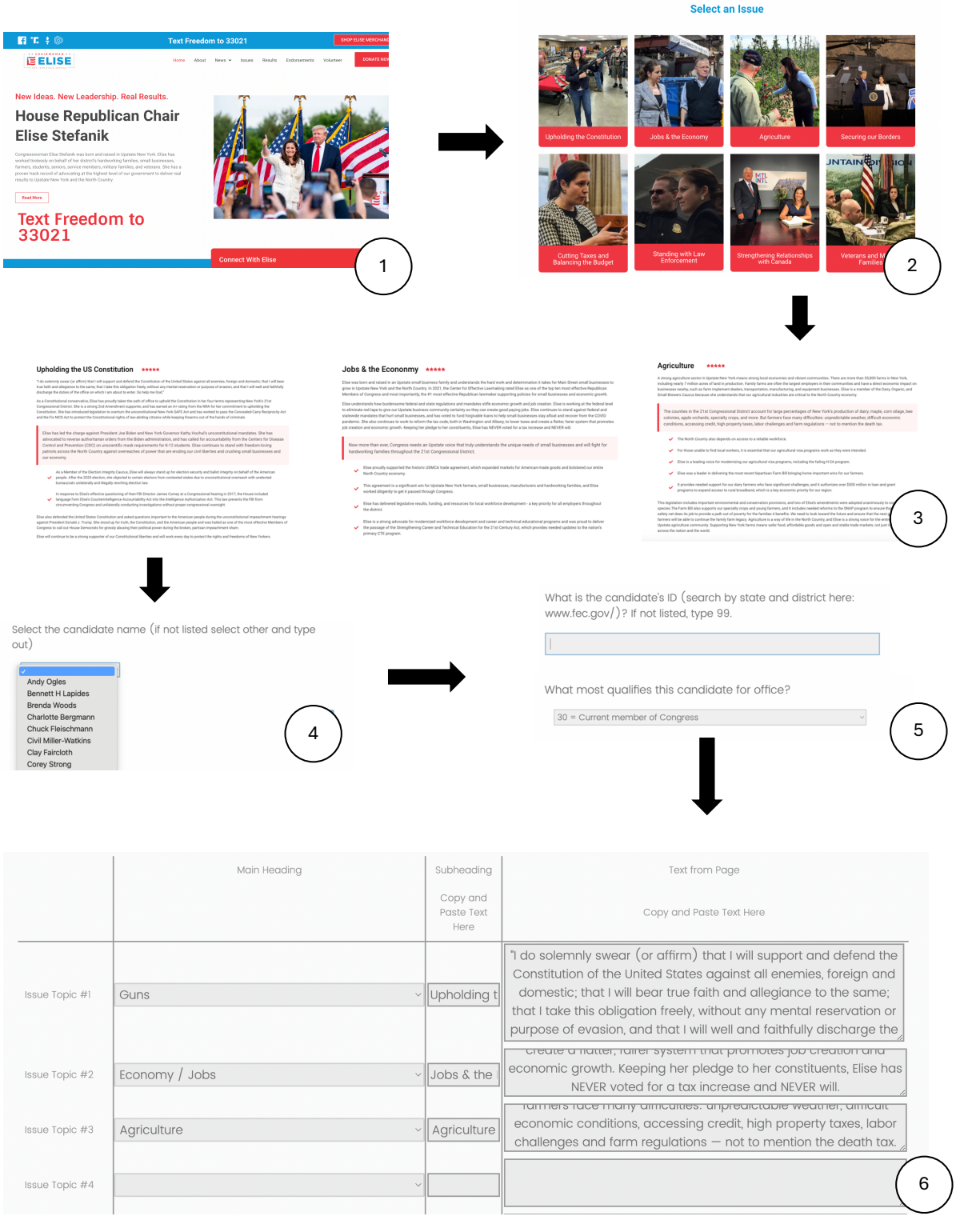
Because of this, the homepage for a given candidate's campaign website may frequently be scraped by an archiving web crawler, but that candidate's campaign platform page is never cataloged or is only cataloged sparsely. For these reasons, we collect all text data in real time.

A small proportion of incumbents—less than 10% of incumbent members of the US who ran for reelection between 2018 and 2022—did not have a campaign website with a platform of policy positions. Most often, these campaign websites would contain only a single landing page that would function to solicit donations. For this minority of incumbents, we collect policy positions stated on their House.gov website. For individuals who do not wish to use scores created with these data, we denote all such observations with a dichotomous indicator in the version of our data available in this paper's replication materials.

Table A1: Predictors for Missingness in Campaign Platforms

	DV: Presence of Policy Platform	
	OLS	GLM
No Incumbent in Election	−0.017	−0.085
	(0.013)	(0.080)
Primary Type: Open	−0.019	−0.124
	(0.012)	(0.075)
Primary Type: Non-Partisan	−0.050*	−0.294*
	(0.017)	(0.102)
Unopposed Primary	−0.045*	−0.278*
	(0.016)	(0.095)
Republican Candidate	−0.027*	−0.173*
	(0.011)	(0.064)
Prior Office-Holder	−0.058*	−0.351*
	(0.017)	(0.108)
Current Incumbent MC	−0.143*	−0.871*
	(0.017)	(0.111)
Logged Pre-Primary Fundraising	0.029*	0.156*
	(0.001)	(0.007)
2020	−0.055*	−0.330*
	(0.013)	(0.078)
2022	−0.012	−0.079
	(0.013)	(0.080)
Constant	0.600*	0.520*
	(0.017)	(0.097)
Observations	6,006	6,006
R^2	0.109	
Log Likelihood		−3,062.394
Note:		*p<0.05

Figure A1: Steps for Campaign Website Text Collection



A2 Left-Right Scaling

A2.1 Testing Units of Aggregation

In the analyses below, we test the performance of CPIs generated using a different unit of text aggregation. Recall that the measurement framework presented in our main paper relies only on paragraphs of text identified by our issue classifier (see discussion in Section A3.3). We use this approach because some candidates discuss multiple issue areas under a single platform point heading (e.g., discussing abortion and health insurance under "Our Healthcare System"). Our goal is to isolate relevant paragraphs of text to limit noise in measurement, a step which follows existing work (e.g., Grimmer and Stewart 2013; Wilkerson and Casas 2017; Laver 2017).

Producing measures at different units of aggregation produces sufficiently different positioning estimates; we demonstrate this by comparing CPI performance to a positioning measure that includes non-relevant text. Our approach for producing this alternative measure follows the same estimation procedure used to generate CPIs, except we estimate positioning based on *all* text under a relevant platform point (e.g., all text under "Our Healthcare System"). We identify issue-relevant platform points as those that contain at least one issue-relevant paragraph identified by our classifier. Hereafter, we refer to this alternative measure as "CPI-Platform-Point." Correlations between CPI and CPI-Platform-Point are listed below; we report within and across-party correlations. Correlations range from relatively high (0.722) to low (0.146). These correlations indicate that text inclusion does impact the scaling of left-right positioning in policy platform text.

- Abortion: Cross-Party (0.563); Dems (0.473); Reps (0.722)
- Education: Cross-Party (0.673); Dems (0.278); Reps (0.721)
- Energy: Cross-Party (0.599); Dems (0.267); Reps (0.503)
- Guns: Cross-Party (0.773); Dems (0.450); Reps (0.672)
- Healthcare: Cross-Party (0.656); Dems (0.296); Reps (0.658)
- Immigration: Cross-Party (0.564); Dems (0.146); Reps (0.516)

The differences in measurement demonstrated above are attributable to text inclusion, given that we alter no other elements of measurement. There is sufficient reason to believe that these differences are specifically attributable to extraneous text inclusion. As we demonstrate in Section A3.3 of the Appendix, our text ensemble stacking classifier has recall and F1 scores of 0.80 or higher, indicating few cases of false negatives (i.e., text which are policy-relevant are not systematically being missed). From this analysis, we conclude that extraneous text inclusion has downstream consequences on measurement that may impact our scaling of policy positioning.

A2.2 Coding Instructions for Paragraph Classification

This task involves reading paragraphs from congressional candidates' campaign platforms and judging whether the text pertains to issue areas of interest. If you believe the paragraph pertains to a given issue area, mark a "1" in that issue area column in the attached Excel spreadsheet. If you do not believe the paragraph pertains to a given issue area, leave that column blank. Paragraph categorizations are not mutually exclusive; this means a single paragraph can be labeled for multiple issues of interest (e.g., one paragraph can pertain to both *Abortion* and *Healthcare*). Reference the descriptions and examples below to determine whether a paragraph pertains to a given issue:

Abortion — the ability to get an abortion
Relevant Examples: pro-life, pro-choice, "reproductive healthcare", Planned Parenthood (if explicitly tied to abortion).
Examples specific to Planned Parenthood:

- We need to defund Planned Parenthood **(yes)**
- We need to defund Planned Parenthood because it provides abortions **(yes)**
- We need to defund Planned Parenthood because they sell fetal body parts **(no)**
- We need to support Planned Parenthood **(yes)**
- We need to support Planned Parenthood because it provides abortions **(yes)**
- We need to support Planned Parenthood because it provides breast cancer screenings **(no)**

Education — access to education, government involvement, school content
Relevant Examples: teacher salaries, school choice, homeschooling, Department of Education, local school boards, free college, college debt, Common Core/No Child Left Behind reform

Energy — resources used to produce energy Relevant Examples: fossil fuels (oil, natural gas, coal), renewables (solar, wind, etc.), "all-of-the-above" policy, energy independence; Does **not** include: discussions solely about climate change, environment, conservation

Guns — the ability to purchase weapons Relevant Examples: gun rights, limits on purchasing, 2nd amendment, red flag laws, concealed carry–requires mention of access/restriction, not just NRA or school shootings alone

Healthcare — access, cost, or quality of healthcare Relevant Examples: Obamacare, insurance, rural health, Medicare, COVID-19, vaccines; abortion included if described as "abortion as healthcare"; Does **not** include: lockdowns or business closures

Immigration — entry to or treatment of immigrants in the U.S.
Relevant Examples: reform, asylum, border security, DACA, services to immigrants, deportation, ICE, family separation

Policing — role, support, or accountability of police Relevant Examples: body cameras, "back the Blue," defunding police, brutality, shootings; Does **not** include: prison reform, school-to-prison pipeline, general criminal justice system

Voting Rights — who can vote and how Relevant Examples: voter ID, ballot fraud, audits, mail-in ballots, holiday voting, ranked-choice voting, enfranchisement (felons, Native Americans); January 6th references must be tied to election security

A2.3 Details on Machine Learning Paragraph Classifier

We train a series of supervised machine learning models to predict whether a campaign platform paragraph discusses topics related to our issues of interest. We train five separate classifiers for each of our six issue areas: decision tree, gradient boosting, logistic regression, support vector machine, and random forest. For all models (other than logistic regression), we select model parameters using a 5-fold cross-validation grid search. We additionally leverage the predictions for all five models to train a logistic regression stacking classifier. Following recommendations from Park and Montgomery (2023), we split our 8,306 labeled campaign platform paragraphs into an 80-20 training-validation split, with 6,602 labeled paragraphs for model training and 1,704

labeled paragraphs for downstream validation. Before model fitting, we pre-process paragraphs by converting all text to lowercase and removing stop words and punctuation. For each classification model (five base models and one ensemble model), we make out-of-sample predictions for all unlabeled paragraphs, as well as the held-out set. Tables A2 to A7 present fit statistics for our machine-learning classifiers. Across all six issue areas of interest, our ensemble classifier produces the highest out-of-sample precision, recall, and F1 Score for our validation data.

Table A2: Abortion: Out-of-Sample Validation Metrics

	Accuracy	Precision	Recall	F1
Logistic	0.975	0.954	0.338	0.500
SVM	0.989	0.940	0.758	0.839
Decision Tree	0.988	0.838	0.838	0.838
Random Forest	0.986	1.000	0.629	0.772
Gradient Boost	0.988	0.890	0.790	0.837
Stacking	0.990	0.924	0.790	0.852

Table A3: Education: Out-of-Sample Validation Metrics

	Accuracy	Precision	Recall	F1
Logistic	0.952	0.981	0.577	0.727
SVM	0.967	0.928	0.759	0.835
Decision Tree	0.957	0.827	0.770	0.797
Random Forest	0.964	0.943	0.716	0.814
Gradient Boost	0.964	0.831	0.844	0.838
Stacking	0.968	0.888	0.812	0.849

Table A4: Energy: Out-of-Sample Validation Metrics

	Accuracy	Precision	Recall	F1
Logistic	0.971	1.000	0.435	0.606
SVM	0.981	0.935	0.682	0.789
Decision Tree	0.981	0.827	0.788	0.807
Random Forest	0.977	0.979	0.564	0.716
Gradient Boost	0.981	0.827	0.788	0.807
Stacking	0.985	0.905	0.788	0.842

Table A5: Guns: Out-of-Sample Validation Metrics

	Accuracy	Precision	Recall	F1
Logistic	0.968	0.950	0.426	0.589
SVM	0.985	0.957	0.752	0.842
Decision Tree	0.985	0.911	0.808	0.857
Random Forest	0.982	0.968	0.685	0.802
Gradient Boost	0.988	0.937	0.842	0.887
Stacking	0.991	0.951	0.876	0.912

Table A6: Healthcare: Out-of-Sample Validation Metrics

	Accuracy	Precision	Recall	F1
Logistic	0.947	0.802	0.503	0.618
SVM	0.965	0.830	0.744	0.785
Decision Tree	0.955	0.721	0.786	0.752
Random Forest	0.953	0.911	0.496	0.642
Gradient Boost	0.963	0.786	0.786	0.786
Stacking	0.971	0.847	0.806	0.826

Table A7: Immigration: Out-of-Sample Validation Metrics

	Accuracy	Precision	Recall	F1
Logistic	0.955	0.948	0.433	0.594
SVM	0.980	0.960	0.771	0.855
Decision Tree	0.978	0.858	0.858	0.858
Random Forest	0.972	0.965	0.653	0.779
Gradient Boost	0.974	0.833	0.826	0.830
Stacking	0.981	0.906	0.842	0.873

A2.4 Full List of Text Pro-Processing Steps

We take the following steps when pre-processing text from our corpus of paragraphs drawn from candidates' campaign platforms:

- All text to lowercase
- String pattern 'medicare for all' to 'medicareforall'
- String pattern 'medicare-for-all' to 'medicareforall'
- String pattern 'pro choice' to 'prochoice'
- String pattern 'pro-choice' to 'prochoice'
- String pattern 'pro-life' to 'prolife'
- String pattern 'de-escalation' to 'deescalation'
- String pattern 'dodd-frank' to 'doddfrank'
- String pattern 'k-12' to 'ktwelve'
- String pattern 'k12' to 'ktwelve'
- String pattern 'pre-k' to 'prek'
- String pattern '4-year' to 'four year'
- String pattern '4 year' to 'four year'
- String pattern '2-year' to 'two year'
- String pattern '2 year' to 'two year'
- String pattern 'non-violent' to 'nonviolent'
- String pattern '2nd amendment' to 'second amendment'
- String pattern '2a' to 'second amendment'
- String pattern 'non-profit' to 'nonprofit'
- String pattern 'non-discrimination' to 'nondiscrimination'
- Replace all '-' with ' '
- Remove all non-alphabetic characters
- Remove extraneous UTC code
- Trim white space

A2.5 Term Selection Strategy & Robustness

To determine the terms that will define the left-most and right-most positions for our eight issue cleavages, we first consulted the policy platforms of far-left and far-right advocacy groups. We additionally consult policy priorities made available by far-left and far-right congressional caucuses. For each issue area, we consulted the following resources:

Abortion: Far-Right

- Eagle Forum: https://eagleforum.org/topics/pro-life.html
- Family Research Council: https://www.frc.org/abortion
- Heritage Foundation: https://www.heritage.org/life-and-family
- National Right to Life: https://nrlc.org/
- Susan B. Anthony: https://sbaprolife.org/about

Abortion: Far-Left

- Center for American Progress: https://www.americanprogress.org/topic/abortion-rights/
- House Progressive Caucus: https://progressives.house.gov/universal-health-care
- Justice Democrats: https://justicedemocrats.com/platform/society/#reproductive-rights
- Our Revolution: https://ourrevolution.com/policy-fights/
- Planned Parenthood: https://www.plannedparenthoodaction.org/issues/abortion
- Reproductive Freedom for All: https://reproductivefreedomforall.org/about/

Education: Far-Right

- Club for Growth: https://www.clubforgrowth.org/issue/education/
- Conservative Caucus: https://www.theconservativecaucus.org/about/stand-for-fight-for
- CPAC: http://ratings.conservative.org/issues?group=B
- Eagle Forum: https://eagleforum.org/topics/education.html
- Family Research Council: https://www.frc.org/education
- FreedomWorks: https://www.freedomworks.org/issue/curriculum/
- Heritage Foundation: https://www.heritage.org/empower-parents-make-education-choices

Education: Far-Left

- Center for American Progress: https://www.americanprogress.org/topic/education-k-12/
- Congressional Progressive Caucus Center: https://www.progressivecaucuscenter.org/debt-free-college
- House Progressive Caucus: https://progressives.house.gov/education
- Justice Democrats: https://justicedemocrats.com/platform/economy/#cancel-student-debt

Energy: Far-Right

- CPAC: http://ratings.conservative.org/issues?group=C
- Conservative Caucus: https://www.theconservativecaucus.org/about/stand-for-fight-for
- FreedomWorks: https://www.freedomworks.org/vote/energy-independence-and-security-act/
- Heritage Foundation: https://www.heritage.org/energy

Energy: Far-Left

- Center for American Progress: https://www.americanprogress.org/topic/clean-energy-2/
- Congressional Progressive Caucus Center: https://www.progressivecaucuscenter.org/issues/environment-climate
- House Progressive Caucus: https://progressives.house.gov/climate-justice
- Our Revolution: https://ourrevolution.com/policy-fights/

Guns: Far-Right

- CPAC: http://ratings.conservative.org/issues?group=A
- Gun Owners of America: https://www.gunowners.org/about-goa/
- National Association for Gun Rights: https://www.nationalgunrights.org/about-us/nagr-pac/
- National Rifle Association: https://home.nra.org/statements/nra-statement-on-gun-control-package/

Guns: Far-Left

- Center for American Progress: https://www.americanprogress.org/topic/gvp/
- Congressional Progressive Caucus Center: https://www.progressivecaucuscenter.org/issues/gun-violence-prevention
- Everytown for Gun Safety: https://www.everytown.org/
- Justice Democrats: https://justicedemocrats.com/platform/society/#gun-safety
- Moms Against Gun Violence: https://momsdemandaction.org/about/

Healthcare: Far-Right

- CPAC: http://ratings.conservative.org/issues?group=F
- Family Research Council: https://www.frc.org/health-care

Healthcare: Far-Left

- Center for American Progress: https://www.americanprogress.org/topic/health-coverage-and-access/
- House Progressive Caucus: https://progressives.house.gov/universal-health-care
- Medicare-For-All Action Network https://actionnetwork.org/letters/medicare-for-all-caucus/
- Our Revolution: https://ourrevolution.com/policy-fights/
- Congressional Progressive Caucus Center: https://www.progressivecaucuscenter.org/medicare-for-all

Immigration: Far-Right

- Conservative Caucus: https://www.theconservativecaucus.org/about/stand-for-fight-for
- Heritage Foundation: https://www.heritage.org/borders-and-crime

Immigration: Far-Left

- Center for American Progress: https://www.americanprogress.org/topic/immigration/
- Congressional Progressive Caucus Center: https://www.progressivecaucuscenter.org/issues/immigrants-rights
- House Progressive Caucus: https://progressives.house.gov/immigrant-rights

We relied on these resources to produce our dictionaries of terms, provided in Table A8 below. In each dictionary, we include at least 25 terms to ensure that our measurement results are not dependent on specific terms. Existing work demonstrates that long-term lists effectively capture the semantic meaning of a latent construct of interest. As a final step, we verify whether our corpus of text accurately reflects the concepts we aim to capture in our positional dictionaries. We achieve this by cross-referencing our dictionaries of terms with the vocabulary of our corpus, and then randomly sampling documents to ensure that the use of terms occurs in expected contexts. An alternative approach to term selection could rely on a computer-assisted, statistical approach to keyword selection. When implementing the approach proposed by King et al. (2017) for computer-assisted keyword discovery, we found our term list more closely reflected our core concepts.

To demonstrate that our estimates are robust to alternative term selections, we randomly sample terms from Table A11 and re-estimate CPIs. We repeat this process a total of 50 times for each issue area. We then correlate each of these scores with the CPIs used in the body of the paper. Consistently, we find very high correlations with our measures, both across and within party. Below are the average correlations across the 50 keyword samples for each issue area. This analysis demonstrates that CPI estimation is not sensitive to excluding or including specific keywords.

- Abortion: Cross-Party (0.983); Dems (0.920); Reps (0.939)
- Education: Cross-Party (0.938); Dems (0.866); Reps (0.919)
- Energy: Cross-Party (0.944); Dems (0.899); Reps (0.933)
- Guns: Cross-Party (0.985); Dems (0.941); Reps (0.967)
- Healthcare: Cross-Party (0.953); Dems (0.900); Reps (0.923)
- Immigration: Cross-Party (0.942); Dems (0.882); Reps (0.885)

Table A8: Issue Cleavages & Full Positional Term Dictionaries

Issue Area	Left-Most Position	Right-Most Position
Abortion	**Pro-Choice**: reproductive, justice, freedom, health, care, healthcare, expand, access, safe, hyde, roe, wade, choice, prochoice, government, codifying, legal, restrictive, decisions, autonomy, control, privacy, doctor, respect, personal, equality, women, sexual, services	**Pro-Life:** life, birth, death, unborn, womb, precious, sanctity, moment, conception, heartbeat, begins, prolife, partial, fetal, born, alive, demand, saving, defund, taxpayer, infanticide, murder, barbaric, innocent, dignity, values, conscientious, immoral, vulnerable, defenseless, ban, outlaw, god, gift, sacred
Education	**Federal Involvement:** public, underfunded, invest, government, responsibility, cuts, investment, expand, teachers, affordable, free, tuition, debt, universal, profit, prek, loan, cancel, college, forgiveness, income, race, live, zip, accessible, equality, pay, salaries, technology, head, start	**Local Control:** parents, bureaucrats, burdensome, tax, state, local, board, know, best, values, control, autonomy, decisions, abolish, eliminate, dismantle, department, decentralize, choice, competition, religious, charter, private, homeschool, vouchers, mandates, core, common, curriculum
Energy	**Renewables Investment:** renewable, clean, green, sustainable, solar, wind, hydroelectric, transition, paris, rejoin, emissions, climate, change, environment, ban, pollution, technology, research, development, incentives, credits, tax, subsidize, funding, investment, new, deal, net, zero, hundred	**Fossil Fuel Investment:** oil, gas, coal, keystone, pipeline, independence, foreign, reliance, deregulate, cut, reopen, restrictions, red, tape, bureaucratic, repeal, free, market, private, competition, companies, china, russia, middle, east, national, security, disruption, unproven, unstable, domestic, expand, production, reserves
Guns	**Increase Restrictions:** mandatory, background, ban, national, hate, registry, database, assault, automatic, ar, ak, military, battlefield, war, waiting, age, years, ghost, loopholes, transfer, traumatic, violence, mass, killing, crisis, epidemic, dealers, manufacturers, trafficking, comprehensive, stricter	**Increase Access:** second, amendment, owner, abiding, right, infringed, inherent, unconstitutional, founding, enumerated, confiscation, unrestricted, repeal, oppose, abolish, any, overturn, enshrined, reciprocity, concealed, carry, hearing, security, freedoms, fundamental, defending, families, protection

Health Insurance	**Publicly Funded:** universal, all, deny, coverage, privilege, guarantee, uninsured, gap, public, medicareforall, singlepayer, free, strengthen, expand, comprehensive, dental, hearing, vision, access, human, right, equality, fundamental, poor, justice, inclusive, profit, preexisting, bankruptcy	**Privatize:** repeal, replace, problems, flawed, free, market, open, competition, freedom, national, state, choice, consumers, individual, patient, provider, insurance, mandate, Washington, burdensome, tape, involvement, bureaucratic, deregulate, tax, hsa, savings, obamacare, socialized
Immigration	**Inclusive:** undocumented, pathway, roadmap, temporary, tps, dreamers, daca, asylum, humane, compassion, dignity, diverse, culture, xenophobic, racist, families, reunification, detention, cages, humanitarian, streamline, backlogs, courts, accessible, opportunity, expand, abolish, ice, education, healthcare	**Exclusive:** enforce, rule, law, secure, build, wall, barrier, surveillance, technology, verify, screening, safety, protect, sanctuary, chain, migration, anchor, birthright, english, skilled, merit, labor, overstay, deport, amnesty, illegal, criminal, terrorists, flooding, porous, close, funding, hire, officers, patrol

A3 Measurement Validation

A3.1 Coding Instructions for Human Judgment Task

See PDF at the end of the document.

A3.2 Coding Descriptive Statistics

Below, we explore descriptive statistics from our human judgment task. Table A9 outlines the coder agreement in categorizing policy platform paragraphs. For most texts, at least two readers agreed on a common score for a given policy platform. Table A10 provides the proportion of documents in each issue area flagged for including only irrelevant text; guidance on coding for irrelevant text varied by issue area and is outlined in greater detail in Section A3.1. On average, readers flagged about 6% or eighteen documents per issue area as including only irrelevant text. Table A11 provides the proportion of documents in each issue area flagged for taking an ambiguous position; guidance on coding for ambiguous positioning text varied by issue area and is outlined in greater detail in Section A3.1. On average, readers flagged about 10% or thirty documents per issue area as including text that was ambiguous in position-taking content .

Table A9: Reader Agreement Rate, By Issue Area

Issue Area	3/3 Agreement	2/3 Agreement	No Agreement
Abortion	0.65	0.33	0.02
Guns	0.52	0.46	0.02
Immigration	0.45	0.53	0.02
Education	0.50	0.43	0.07
Energy	0.56	0.39	0.05
Healthcare	0.49	0.43	0.08

Table A10: Flagged Irrelevant Text, by Issue Area

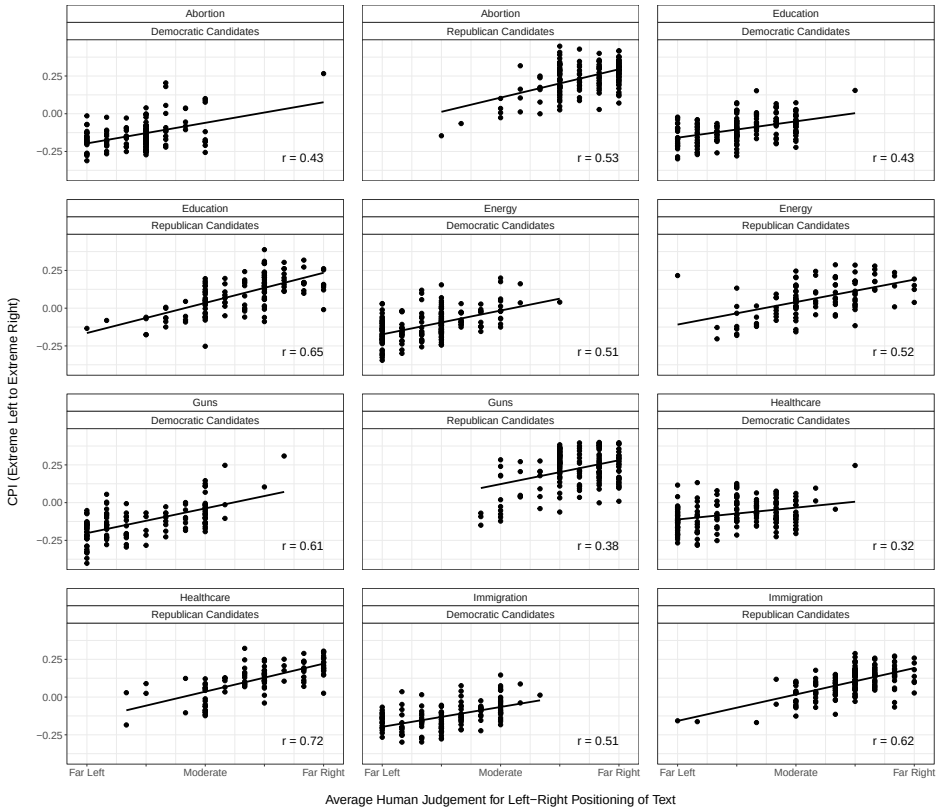
Issue Area	Reader A (PI)	Reader B (Student)	Reader C (Student)	Reader D (Student)
Abortion	0.01	0.07	0.03	—
Guns	0.01	—	0.01	0.01
Immigration	0.02	0.05	0.03	—
Education	0.09	0.08	—	0.16
Energy	0.06	0.04	—	0.04
Healthcare	0.01	0.03	0.00	—

Table A11: Flagged Ambiguous Text, by Issue Area

Issue Area	Reader A (PI)	Reader B (Student)	Reader C (Student)	Reader D (Student)
Abortion	0.02	0.02	0.03	—
Guns	0.03	—	0.04	0.06
Immigration	0.08	0.10	0.05	—
Education	0.05	0.16	—	0.10
Energy	0.05	0.12	—	0.05
Healthcare	0.02	0.10	0.06	—

A3.3 CPI-Human Judgment Correlations by Party

Figure A2: CPI Validation with Human Judgments



A3.4 CPI-PAC Giving: Model Outputs

Table A12: CPI Validation with Pro-Choice PAC Giving: Democratic Candidates

	Dependent Variable: Left-Leaning Abortion PAC Contribution			
	(Main Paper Model)	(Bivariate)	(Alternative 1)	(Alternative 2)
CPI-Abortion	−0.458*	−0.482*		
	(0.127)	(0.124)		
CPI-Abortion (Imputed at party mean)			−0.493*	
			(0.123)	
CPI-Abortion (Imputed at moderate)				−0.434*
				(0.059)
DIME-CFscore	3.259*		1.937*	2.207*
	(0.631)		(0.251)	(0.258)
Constant	2.973*	−0.178	1.125*	2.012*
	(0.667)	(0.259)	(0.333)	(0.236)
Observations	470	477	1,238	1,238

Note: CPIs and CFscores range from negative (left) to positive (right). CFscore coefficients here are positive, indicating that as a candidate becomes more right-leaning, they are *more likely* to receive pro-choice PAC funding. Running a bivariate model with CFscore and PAC giving produces the same counter-intuitive result. $p < 0.05$.

Table A13: CPI Validation with Pro-Life PAC Giving: Republican Candidates

	Dependent Variable: Right-Leaning Abortion PAC Contribution			
	(Main Paper Model)	(Bivariate)	(Alternative 1)	(Alternative 2)
CPI-Abortion	0.253*	0.264*		
	(0.126)	(0.114)		
CPI-Abortion (imputed at party mean)			0.213*	
			(0.108)	
CPI-Abortion (imputed at moderate)				0.535*
				(0.070)
DIME-CFscore	−2.689*		−0.267	−0.475*
	(0.659)		(0.205)	(0.210)
Constant	4.213*	0.666*	0.461	0.753*
	(0.860)	(0.224)	(0.308)	(0.256)
Observations	434	451	1,115	1,115

Note: CPIs and CFscores range from negative (left) to positive (right). CFscore coefficients here are negative, indicating that as a candidate becomes more right-leaning, they are *less likely* to receive pro-choice PAC funding. Running a bivariate model with CFscore and PAC giving produces the same counter-intuitive result. $p < 0.05$.

Table A14: CPI Validation with Clean Energy PAC Giving: Democratic Candidates

	Dependent Variable: Left-Leaning Energy PAC Contribution			
	(Main Paper Model)	(Bivariate)	(Alternative 1)	(Alternative 2)
CPI-Energy	−0.166	−0.087		
	(0.116)	(0.112)		
CPI-Energy (Imputed at party mean)			−0.163	
			(0.123)	
CPI-Energy (Imputed at moderate)				−0.304*
				(0.108)
DIME-CFscore	2.395*		1.948*	2.014*
	(0.364)		(0.258)	(0.260)
Constant	0.453	−1.523*	−0.153	−0.115
	(0.307)	(0.120)	(0.221)	(0.214)
Observations	706	716	1,238	1,238

Note: CPIs and CFscores range from negative (left) to positive (right). CFscore coefficients here are positive, indicating that as a candidate becomes more right-leaning, they are *more likely* to receive left-leaning energy PAC funding. Running a bivariate model with CFscore and PAC giving produces the same counter-intuitive result. $p < 0.05$.

Table A15: CPI Validation with Fossil Fuels PAC Giving: Republican Candidates

	Dependent Variable: Right-Leaning Energy PAC Contribution			
	(Main Paper Model)	(Bivariate)	(Alternative 1)	(Alternative 2)
CPI-Energy	0.363*	0.319*		
	(0.127)	(0.115)		
CPI-Energy (imputed at party mean)			0.326*	
			(0.120)	
CPI-Energy (imputed at moderate)				0.490*
				(0.089)
DIME-CFscore	−2.358*		−1.554*	−1.498*
	(0.527)		(0.240)	(0.239)
Constant	2.918*	0.042	1.499*	1.638*
	(0.641)	(0.164)	(0.318)	(0.295)
Observations	339	359	1,115	1,115

Note: CPIs and CFscores range from negative (left) to positive (right). CFscore coefficients here are negative, indicating that as a candidate becomes more right-leaning, they are *less likely* to receive right-leaning energy PAC funding. Running a bivariate model with CFscore and PAC giving produces the same counter-intuitive result. $p < 0.05$.

Table A16: CPI Validation with Pro-Gun Control PAC Giving: Democratic Candidates

	Dependent Variable: Left-Leaning Gun PAC Contribution			
	(Main Paper Model)	(Bivariate)	(Alternative 1)	(Alternative 2)
CPI-Guns	−0.284*	−0.305*		
	(0.091)	(0.090)		
CPI-Guns (imputed at party mean)			−0.301*	
			(0.094)	
CPI-Guns (imputed at moderate)				−0.222*
				(0.053)
DIME-CFscore	0.689		0.841*	0.966*
	(0.379)		(0.212)	(0.215)
Constant	−0.358	−1.041*	−0.440	0.040
	(0.398)	(0.194)	(0.258)	(0.190)
Observations	538	546	1,238	1,238

Note: CPIs and CFscores range from negative (left) to positive (right). CFscore coefficients here are positive, indicating that as a candidate becomes more right-leaning, they are *more likely* to receive pro-gun control PAC funding. Running a bivariate model with CFscore and PAC giving produces the same counter-intuitive result. $p < 0.05$.

Table A17: CPI Validation with Pro-Gun Rights PAC Giving: Republican Candidates

	Dependent Variable: Right-Leaning Gun PAC Contribution			
	(Main Paper Model)	(Bivariate)	(Alternative 1)	(Alternative 2)
CPI-Guns	0.456*	0.411*		
	(0.087)	(0.077)		
CPI-Guns (imputed at party mean)			0.400*	
			(0.078)	
CPI-Guns (imputed at moderate)				0.491*
				(0.060)
DIME-CFscore	−3.610*		−0.821*	−0.983*
	(0.561)		(0.231)	(0.234)
Constant	4.872*	0.257	0.904*	1.380*
	(0.722)	(0.150)	(0.303)	(0.286)
Observations	555	530	1,115	1,115

Note: CPIs and CFscores range from negative (left) to positive (right). CFscore coefficients here are negative, indicating that as a candidate becomes more right-leaning, they are *less likely* to receive pro-gun rights PAC funding. Running a bivariate model with CFscore and PAC giving produces the same counter-intuitive result. $p < 0.05$.

A4 Main Paper Results

A4.1 Establishing Heterogeneity in Policy-Level Positioning

Figure A3: Raw CPI Pairwise Correlations, By Candidate Type

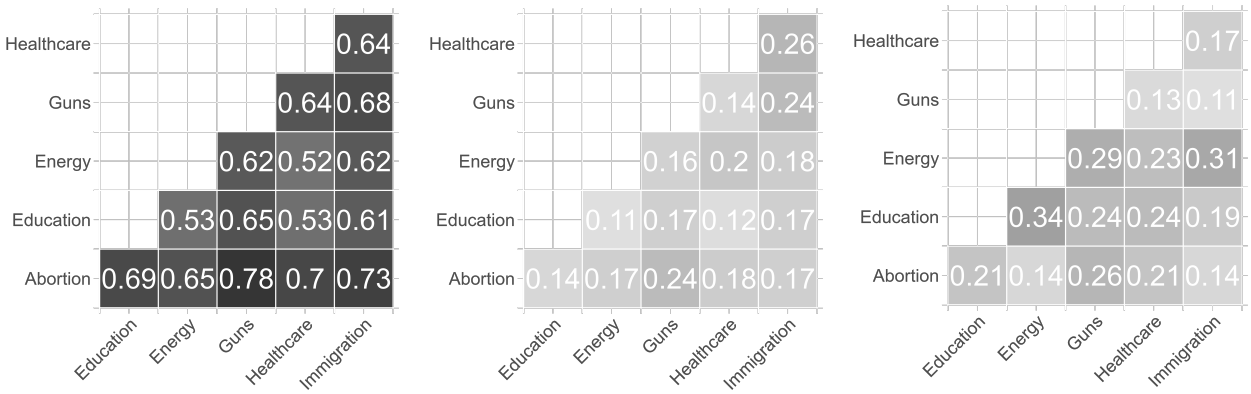


Table A18: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Abortion

Percentile	Candidate (Year)	Excerpt from Campaign Platform
10th (L)	Mike Quigley (D-IL, 2018)	"Mike trusts women to make their own healthcare choices and is a stalwart champion for women's reproductive rights...He supports a full repeal of the Hyde Amendment and was active in the fight against anti-choice measures being included in the Affordable Care Act."
30th	Stephen Lynch (D-MA, 2020)	"Stephen supports the Equal Rights Amendment... When Republicans tried to defund women's health organizations, like Planned Parenthood, Stephen stood on the floor of the House of Representatives to fight the legislation."
50th	Mark Meadows (R-NC, 2018)	"As a longtime pro-life activist, I will work in Congress to end the tragedy of abortion. I'll work just as hard to reduce barriers to adoption, and promote policies that support groups offering hope and assistance to mothers facing crisis pregnancies."
70th	Sam Graves (R-MO, 2022)	"I have supported various bills and policies to defund Planned Parenthood, and will also do whatever possible to defend life. I believe that we should protect the lives of the unborn under all but the direst of circumstances."
90th (R)	Bill Huizenga (R-MI, 2022)	"Protecting life beginning with conception is a value that cannot be compromised... I stressed this for 10 years as the Right to Life representative for my church in Zeeland and continue to fight for the unborn as a member of Congress."

Table A19: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Education

Percentile	Candidate (Year)	Excerpt from Campaign Platform
10th (L)	Donald McEachin (D-VA, 2022)	"Donald will work towards universal access to pre-K education and affordable, high-quality childcare options. He will also seek overdue investments in our public schools, ensuring that teachers have the resources and support they need to close achievement gaps and best help our children."
30th	Kathy Castor (R-FL, 2018)	"Kathy believes that every child deserves a good education and she has worked hard to ensure that each student has the chance to succeed through her work with Pell Grants, student loans, Upward Bound and Head Start. She has also worked diligently to fend off Republican cuts to students, schools..."
50th	Daniel Crenshaw (R-TX, 2020)	"It is crucial that we prepare the next generation for the workforce. We must promote vocational training as a sensible alternative to a college education and allow for Pell Grant flexibility to allow students that option..."
70th	Ben Cline (R-VA, 2022)	"I will work to ensure we put children first, not bureaucrats and unions...it is essential that we empower parents to have a leading role in directing their children's education."
90th (R)	Ron Estes (R-TN, 2022)	"I will never support the Federal Government taking power further away from We the People. From our second amendment rights, to the education of our children, to our right to worship as we choose, our liberties must be protected from those who would...tell us how we educate our children."

Table A20: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Energy

Percentile	Candidate (Year)	Excerpt from Campaign Platform
10th (L)	Seth Moulton (D-MA, 2020)	"The Green New Deal needs to be built around green jobs and clean energy, decarbonization and breakthrough technologies, and access to green energy for the developing world—and America should lead the way in winning this moral and economic opportunity..."
30th	Paul Tonko (D-NY, 2022)	"I am proud to sit on the Committee for Science, Space and Technology, and the Energy and Commerce Committee. As an engineer, I understand the importance of clean energy and the continued advancement of innovation in technology sectors."
50th	Tom Reed (R-NY, 2018)	"Tom believes that low-cost energy is another key driver of job creation and economic growth. It's why he supports an all-of-the-above approach to our energy needs to improve our nation's energy infrastructure, incentivize energy exploration...promote the workforce of tomorrow."
70th	Dana Rohrabacher (R-CA, 2018)	"Dana has strongly supported efforts to make America energy self-sufficient. He stood up for expanding domestic oil and gas production, be it by fracking or other approaches. He also championed the development of new safe nuclear alternatives...fought to open federal lands to solar power projects."
90th (R)	Virginia Foxx (R-NC, 2020)	"God blessed America with abundant resources, including fossil fuels and renewable energy sources. All of those resources should be safely developed and Washington bureaucrats should not be able to arbitrarily lock away those resources on public land. Whenever possible, the federal government should remove red-tape barriers to energy development..."

Table A21: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Guns

Percentile	Candidate (Year)	Excerpt from Campaign Platform
10th (L)	Alma Adams (D-NC, 2020)	"Gun control legislation is essential to fighting crime, protecting our citizens, and reducing dangers our law enforcement face each day. Congress needs to ban certain high powered assault weapons, close the gun show loophole, require background checks for purchases of firearms, and improve the tracking of firearms in our country."
30th	John Garamendi (D-CA, 2020)	"Garamendi also supports stopping people on terrorist watch lists from purchasing arms, particularly guns designed to kill as many people as possible as quickly as possible."
50th	Lloyd Smucker (R-PA, 2020)	"Voted for concealed carry reciprocity and NRA-supported improvements to background checks. Cosponsored Federal Firearms Licensee Protection Act (H.R. 3790), which would enhance penalties for the theft of a firearm from federally licensed firearm stores..."
70th	Brian Babin (R-TX, 2022)	"As a 30-year member of the NRA, I am a rock-solid defender of the Second Amendment...In Congress, I have actively introduced, cosponsored, and voted for important legislation that protects the gun rights of all law abiding American citizens."
90th (R)	Doug LaMalfa (R-CA, 2020)	"I am an ardent defender of our Second Amendment rights. I believe firmly in the individual right to keep and bear arms. I will fight any attempt to water down or weaken our rights, while pushing to ensure our rights to hunt and defend ourselves are expanded. I am proud to have earned an A rating and the endorsements of the NRA and Gun Owners of America."

Table A22: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Healthcare

Percentile	Candidate (Year)	Excerpt from Campaign Platform
10th (L)	Katherine Clark (D-CT, 2022)	"We believe access to health care should be a right, not a privilege. Katherine Clark is: An original co-sponsor of Medicare for All; expanded access to affordable care through the American Rescue Plan."
30th	Frank Pallone (D-NJ, 2022)	"Despite the consistent efforts of the Republican majority in Congress to deny affordable health care to millions of Americans and drive up the cost of care for everyone, I remain vigilant to protect the right to health care for all. I will continue to fight to fund the Children's Health Insurance Program and fight against any effort to privatize Medicare for our seniors."
50th	John Rutherford (R-FL, 2022)	"Fought for lower treatment and premium costs across the board; protected pre-existing condition coverage; fought for dependents to stay on parents' insurance until age 26; supported President Trump's efforts to lower prescription drug prices for seniors."
70th	Don Bacon (R-NE, 2022	"I will continue to work towards market-based options to bring down healthcare costs...We can increase options and competition by encouraging associational pools and allowing insurance companies to compete across state lines...I support protecting coverage for those with pre-existing conditions."
90th (R)	David Kustoff (R-TN, 2022)	"I am proud we repealed the Obamacare Individual Mandate Tax. This means you are no longer forced by the government to buy insurance you don't want or need. I will continue to work to dismantle Obamacare once and for all so we can offer the best healthcare in the world."

Table A23: Campaign Platform Excerpts at 20 Percentile-Rank Intervals for CPI-Immigration

Percentile	Candidate (Year)	Excerpt from Campaign Platform
10th (L)	Ritchie Torres (D-NY, 2022)	"Ritchie is ready to fight for an agenda that lifts up the millions of immigrants who come to our country seeking a better life...Not only do we need a pathway to citizenship for the millions of undocumented individuals, but Ritchie also believes we must make our health care, educational, and vocational programs more inclusive..."
30th	Ed Case (D-HI, 2022)	"Support expanded legal immigration...Address today's problems today and without handing them off to our children."
50th	Ron Estes (R-KS, 2022)	"For decades, Washington insiders have ignored the realities on our Southern border. This is unacceptable...Our immigration process must deter criminals and terrorists from entering illegally while providing a fair, efficient process for law abiding individuals wishing to pursue the American dream."
70th	Gary Palmer (R-AL, 2018)	"I am opposed to amnesty for illegal immigrants. I signed the FAIR Congressional Task Force's No-Amnesty pledge to 'oppose legislation that would grant any form of work authorization to illegal aliens.' The first steps to any reform of America's immigration system will have to be securing our borders and enforcing current immigration law."
90th (R)	Jeff Duncan (R-SC, 2020)	"You will not find anyone in Congress who supports securing our borders, and enforcing our immigration laws more than Jeff Duncan. Jeff has stood with President Trump as he has fought against illegal immigration...has been a long advocate for banning sanctuary cities, requiring mandatory E-Verify, eliminating birthright citizenship, and ending chain migration."

Table A24: Word Exclusivity Comparison: CPI-Education

Percentile Comparison	**Higher Percentile** *More Right-Leaning*	**Lower Percentile** *More Left-Leaning*
90th vs. 70th	parents, federal, government, common, choice, local, core, children, department, control, best, washington, parental, states, theory	college, higher, students, debt, universities, degree, quality, veterans, jobs, student, good, democracy, loans, high, loan
70th vs. 50th	parents, choice, children, race, theory, school, crt, schools, core, history, choose, results, failing, state, system	women, access, control, decision, health, reproductive, woman, decisions, may, trimester, court, state, doctor, options, personal
50th vs. 30th	think, digital, questions, foster, matters, material, empowerment, classrooms, senate, victims, promoting, social, youth, entrepreneurship, may	college, loans, debt, programs, loan, infrastructure, community, tuition, economy, program, benefits, investments, student, stem
30th vs. 10th	david, mike, policies, allows, students, money, assault, county, states, mental, lgbtq, online, it's, congresswoman, focused	college, higher, students, debt, universities, degree, quality, veterans, jobs, student, good, democracy, loans, high, loan

Table A25: Word Exclusivity Comparison: CPI-Energy

Percentile Comparison	**Higher Percentile** *More Right-Leaning*	**Lower Percentile** *More Left-Leaning*
90th vs. 70th	oil, independence, pipeline, keystone, gas, us, president, production, foreign, free, america, barrels, american, prices, countries	clean, fossil, research, future, change, power, solar, environment, air, district, investing, planet, connecticut, seniors, wind
70th vs. 50th	biden, prices, nuclear, domestic, reliable, energy, connecticut, seniors, home, natural	housing, clean, climate, iowa, sustainable, part, jobs, homes, invest, green
50th vs. 30th	oil, fracking, natural, economics, gas, knowledge, hydrogen, term, fighting, technological, example, mike, near, sensible, market	climate, renewable, change, fossil, fuel, carbon, technology, investments, public, communities, efficiency, fracking, economics, large
30th vs. 10th	coal, foreign, california, people, it's, nuclear, oil, sense, believe, years, based, allow, average, makes, tens	clean, 2050, helped, fighting, reach, action, cosponsored, crisis, transition, act, ensure, massachusetts, congress, bold, climate

Table A26: Word Exclusivity Comparison: CPI-Guns

Percentile Comparison	**Higher Percentile** *More Right-Leaning*	**Lower Percentile** *More Left-Leaning*
90th vs. 70th	right, rights, liberties, tom, firmly, fundamental, arms, bear, constitution, protect, protected, self, tirelessly, liberty, amendment	guns, laws, firearms, criminals, violence, ban, national, also, crime, firearm, use, hands, time, preserving, armed
70th vs. 50th	amendment, right, second, bear, fight, defend, shall, arms, infringed, rights, member, conservative, keep, constitutional, supporter	violence, background, mass, checks, gun, safety, shootings, common, sense, weapons, need, can, schools, school, public
50th vs. 30th	right, arms, bear, rights, amendment, constitution, second, 2nd, citizens, government, constitutional, defend, infringed, free, process	violence, gun, ban, background, loophole, universal, domestic, prevention, checks, weapons, health, assault, closing, hands, action
30th vs. 10th	amendment, right, second, rights, bear, arms, carry, 2nd, law-abiding, concealed, citizens, court, sensible, handgun, constitutional	assault, weapons, violence, background, close, high-capacity, magazines, ban, checks, hr, charleston, streets, deadly, domestic, federal

Table A27: Word Exclusivity Comparison: CPI-Healthcare

Percentile Comparison	**Higher Percentile** *More Right-Leaning*	**Lower Percentile** *More Left-Leaning*
90th vs. 70th	obamacare, repeal, government, competition, lines, market, choices, free-market, replace, patient-centered, mandate, healthcare, o, free, obama	healthcare, right, abortion, gender, access, dental, reproductive, universal, human, 2019, housing, guarantee, standard, regardless, medical
70th vs. 50th	government, bureaucrats, obamacare, competition, market, patient, maine, dan, state, kentucky, health-care, choice, accounts, oklahoma, solutions	massachusetts, fight, reproductive, access, brought, addiction, ensure, lgbtq, communities, care, coronavirus, support, discrimination, david, black
50th vs. 30th	costs, insurance, fraud, taxes, h.r, premiums, business, marketplace, consumers, dr, voted, corporations, individual, plan, want	dr, mental, women, act, human, universal, health, 2015, public, coverage, violence, drug, essential, h.r, planned
30th vs. 10th	aca, option, competition, republicans, plans, market, bipartisan, cost, congress, already, lifetime, legislature, enrollment, get, now	medicare, security, prescription, reproductive, family, social, women's, public, women, everyone, sick, privatize, coverage, committee, seniors

Table A28: Word Exclusivity Comparison: CPI-Immigration

Percentile Comparison	**Higher Percentile** *More Right-Leaning*	**Lower Percentile** *More Left-Leaning*
90th vs. 70th	border, wall, southern, patrol, build, chain, across, security, dangerous, books, enforce, flood, stop, cities, finish	reform, citizenship, path, nations, better, towards, tom, system, paso, families, ago, immigrate, undocumented, opportunities, alien
70th vs. 50th	illegal, borders, sanctuary, secure, cities, aliens, end, tom, securing, border, president, barrier, biden, city, crisis	u.s, home, children, dream, person, get, i.c.e, citizenship, comprehensive, individuals, family, undocumented, dreamers, age, path
50th vs. 30th	illegal, border, laws, illegally, amnesty, legally, aliens, legal, security, southern, entering, wall, threats, problem, secure	families, comprehensive, communities, dreamers, detention, undocumented, act, pathway, reform, daca, immigrants, economy, citizenship, community, ice
30th vs. 10th	border, security, borders, illegally, dangerous, can, means, productive, felony, tax, need, business, employers, require, set	immigrant, detention, dreamers, advocate, families, family, tps, fight, asylum, permanent, believes, centers, child, treatment, end

A4.2 District Policy Opinion & Candidate Policy-Level Positioning

Table A29: Relationship Between CPI and District Policy-Specific Opinion, Full Models

	Dependent variable: Candidate Positioning Index (CPI)					
	Abortion	Education	Energy	Guns	Healthcare	Immigration
District Opinion	0.048***	0.027**	0.067***	0.070***	0.062***	0.080***
	(0.006)	(0.013)	(0.009)	(0.006)	(0.007)	(0.007)
Republican	1.692***	1.397***	1.323***	1.554***	1.385***	1.405***
	(0.022)	(0.026)	(0.035)	(0.025)	(0.026)	(0.026)
Experienced Challenger	−0.019	0.091**	−0.025	0.047	−0.027	−0.013
	(0.029)	(0.039)	(0.049)	(0.035)	(0.038)	(0.036)
Incumbent	−0.044	−0.041	0.061	0.063**	0.210***	0.039
	(0.026)	(0.034)	(0.039)	(0.030)	(0.031)	(0.032)
Female	−0.035*	0.006	−0.031	0.008	−0.140***	−0.104***
	(0.021)	(0.024)	(0.032)	(0.024)	(0.024)	(0.024)
Open Seat	−0.032	−0.057*	−0.075*	−0.015	−0.024	0.032
	(0.024)	(0.031)	(0.039)	(0.028)	(0.030)	(0.030)
2020	−0.022	−0.174***	−0.201***	−0.062**	−0.108***	−0.054*
	(0.025)	(0.031)	(0.039)	(0.028)	(0.028)	(0.030)
2022	0.029	−0.174***	−0.129***	−0.010	−0.177***	0.018
	(0.025)	(0.031)	(0.038)	(0.028)	(0.031)	(0.030)
Constant	−0.861***	−0.462***	−0.337***	−0.846***	−0.443***	−0.762***
	(0.024)	(0.028)	(0.033)	(0.026)	(0.026)	(0.028)
Observations	2,162	3,311	2,282	2,556	3,174	2,931

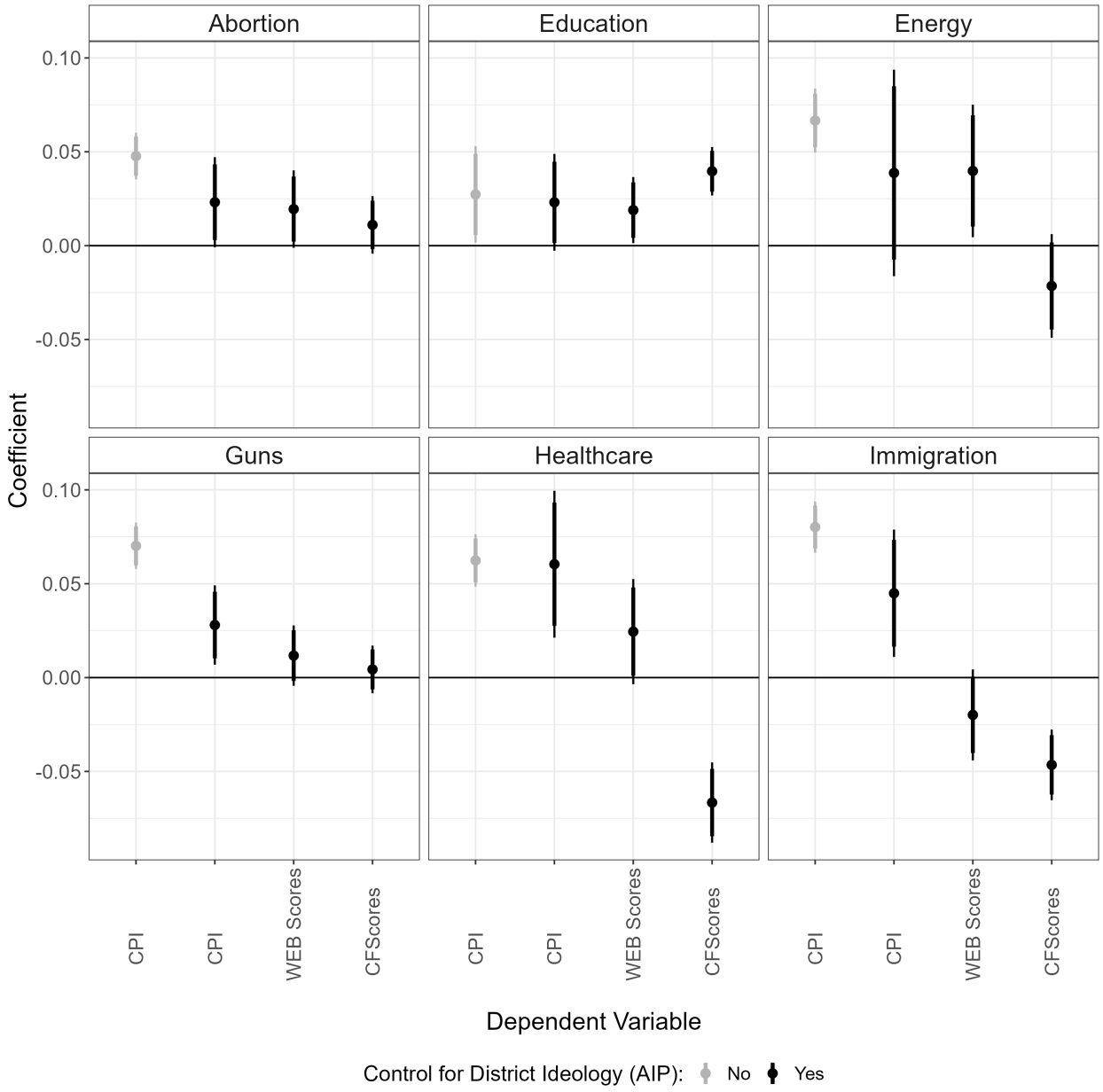
Note: *p<0.1; **p<0.05; ***p<0.01

Table A30: Relationship Between CPI and District Policy-Specific Opinion, Controlling for District Ideological Orientation

	Dependent variable: Candidate Positioning Index (CPI)					
	Abortion	Education	Energy	Guns	Healthcare	Immigration
District Opinion	0.023*	0.023*	0.039	0.028***	0.060***	0.045***
	(0.012)	(0.013)	(0.028)	(0.011)	(0.020)	(0.017)
District Ideology (AIP)	0.287**	0.340***	0.312	0.583***	0.031	0.391**
	(0.123)	(0.074)	(0.299)	(0.117)	(0.200)	(0.178)
Republican	1.684***	1.383***	1.320***	1.532***	1.378***	1.398***
	(0.023)	(0.027)	(0.035)	(0.025)	(0.026)	(0.026)
Experienced Challenger	−0.019	0.093**	−0.019	0.055	−0.021	−0.008
	(0.029)	(0.039)	(0.049)	(0.035)	(0.038)	(0.037)
Incumbent	−0.047*	−0.030	0.061	0.065**	0.213***	0.041
	(0.027)	(0.034)	(0.039)	(0.030)	(0.031)	(0.032)
Female	−0.037*	0.011	−0.031	0.011	−0.143***	−0.103***
	(0.021)	(0.024)	(0.032)	(0.024)	(0.024)	(0.025)
Open Seat	−0.034	−0.062**	−0.069*	−0.016	−0.022	0.029
	(0.024)	(0.031)	(0.039)	(0.028)	(0.030)	(0.030)
2020	−0.018	−0.162***	−0.211***	−0.051*	−0.110***	−0.053*
	(0.025)	(0.032)	(0.039)	(0.028)	(0.028)	(0.030)
2022	0.036	−0.157***	−0.127***	0.007	−0.178***	0.016
	(0.025)	(0.031)	(0.038)	(0.028)	(0.031)	(0.030)
Constant	−0.854***	−0.455***	−0.330***	−0.836***	−0.441***	−0.750***
	(0.024)	(0.028)	(0.035)	(0.026)	(0.027)	(0.029)
Observations	2,144	3,284	2,263	2,538	3,148	2,911

Note: *p<0.1; **p<0.05; ***p<0.01

Figure A4: OLS Regression Model Coefficients: District Policy-Level Opinion and Candidate Positioning



Note: Panels display model coefficients across various specifications. The far-left coefficients are drawn from base models (Appendix Table A29). The middle-left coefficients are drawn from models replicating these analyses, controlling for district overall ideological orientation (Appendix Table A30). Middle- and far-right coefficients fit the same models, with alternative dependent variables measuring candidates' overall ideological positioning from campaign website data (WEB Scores; Case 2025) and donation data (CFScores, Bonica 2024). Bars indicate 95% and 90% confidence intervals.

Table A31

	Dependent variable:	
	Variance in CPIs	
	(1)	(2)
Same Party Vote	-0.0002^{***}	
	(0.00004)	
Same Party Ideology		-0.017^{***}
		(0.003)
Partisan Primary	-0.003^{*}	-0.002
	(0.002)	(0.002)
Fundraising (Logged)	-0.0004^{***}	-0.0004^{***}
	(0.0001)	(0.0001)
Republican	-0.005^{***}	-0.005^{***}
	(0.001)	(0.001)
Open Seat	-0.001	-0.001
	(0.001)	(0.001)
Female	0.002^{*}	0.002^{*}
	(0.001)	(0.001)
Experienced Challenger	-0.0001	-0.0005
	(0.002)	(0.002)
Incumbent	0.002	0.002
	(0.002)	(0.002)
2020 Election Year	-0.002	-0.003^{**}
	(0.001)	(0.001)
2022 Election Year	-0.003^{**}	-0.003^{**}
	(0.001)	(0.001)
Constant	0.053^{***}	0.045^{***}
	(0.003)	(0.002)
Observations	3,965	3,965
Note:	*p<0.1; **p<0.05; ***p<0.01	

References

Bonica, Adam. 2024. "Database on Ideology, Money in Politics, and Elections: Public version 4.0 [Computer file].".

Case, Colin R. 2025. "Measuring Strategic Positioning in Congressional Elections." *Journal of Politics* .

Grimmer, Justin, and Brandon M Stewart. 2013. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political analysis* 21(3): 267–297.

King, Gary, Patrick Lam, and Margaret E Roberts. 2017. "Computer-assisted keyword and document set discovery from unstructured text." *American Journal of Political Science* 61(4): 971–988.

Laver, Michael. 2017. "Measuring policy positions in political space." *Annual Review of Political Science* 17(1): 207–223.

Park, Ju Yeon, and Jacob M. Montgomery. 2023. "Towards a framework for creating trustworthy measures with supervised machine learning.".

Wilkerson, John, and Andreu Casas. 2017. "Large-scale computerized text analysis in political science: Opportunities and challenges." *Annual Review of Political Science* 20(1): 529–544.

Coding Abortion Policy Statements (Pro-Choice vs. Pro-Life)

This task involves reading text from congressional candidates' campaign platforms and judging how much they are left-leaning (very left, left) or right-leaning (very right, right). If you believe the text expresses a centrist position, select the "centrist" option. If you believe the text does not express a clear position, select the "ambiguous" option. If you believe the text does not express a position relevant to abortion policy, select the "irrelevant" option.

What are "left" or "right" abortion policies?

Very left positions tend to advocate for the following:

- Access to abortion without restrictions, removing any and all obstacles
- Safeguarding abortion rights through **federal legislation**
    - Codifying Roe vs. Wade
    - Repeal the Hyde Amendment, global gag rule
- Views abortion as a **fundamental right of women**
    - "Abortions are healthcare", "reproductive justice", "reproductive rights"

Left positions tend to advocate for the following:

- **General pro-choice rhetoric** without explicit discussions of putting protections into law
- Views abortion as a choice that should be available to women
    - "Pro-choice", "Protecting women's right to choose"
- May express that they personally do not believe in abortion, but still should be available

Right positions tend to advocate for the following:

- **General pro-life rhetoric**, abortions being unlawful in nearly all cases
    - Exceptions include **rape, incest, child pregnancy**
- Simply states they are "pro-life", "right to life", or "life is precious"

Very right positions tend to advocate for the following:

- Explicitly says **no abortions** under any circumstances **(only exception is the life of the mother)**
- Views abortion as murder/homicide
- Outlawing abortion through federal legislation
- Includes rhetoric like: **"Life begins at conception"**

What is a centrist abortion policy?

- Abortions **available to the general public with restrictions** (limit to first trimester, **20 weeks**)
- Takes an equivalent proportion of left and right positions

What is an ambiguous abortion policy?

- Discusses abortion in terms of reproductive education, and adoption options without taking an explicit pro-life/pro-choice position on the issue
- Discusses reproductive healthcare without reference to bodily choices

What is an irrelevant abortion policy?

- Discusses reproduction (e.g., access to contraception, education), women's healthcare, or Planned Parenthood without any discussion of abortion

Coding Education Policy Statements (Federal vs. Local)

This task involves reading text from congressional candidates' campaign platforms and judging how much they are left-leaning (very left, left) or right-leaning (very right, right). If you believe the text expresses a centrist position, select the "centrist" option. If you believe the paragraph does not express a clear position, select the "ambiguous" option. If you believe the text does not express a position relevant to immigration policy, select the "irrelevant" option.

What are "left" or "right" education policies?

Very left positions tend to advocate for the following:

- Education is made better *with* ***maximal government intervention***
- Left positions *plus some combination of:*
    - **Universal pre-k, free college for all, college for all**
    - **Public college free, community college free**

Left positions tend to advocate for the following:

- Education is made better *with* ***more government intervention***
    - Prioritizing **funding for public schools**, tax incentives for vocational programs, making **college more affordable**, providing some loan forgiveness/**better lending programs,** giving **public school teachers the resources**/pay they need, increasing federal funding for schools, creating more **equity across public schools, reducing the cost of college**

Right positions tend to advocate for the following:

- Education is made better *with* ***less government intervention***
    - Curriculum choice, **parent choice over schools**, homeschooling, more options beyond public school, remove Common Core, No Child Left Behind

Very right positions tend to advocate for the following:

- Education is made better *with* ***minimal government intervention***
- Right positions *plus some combination of:*
    - **Disband/weaken the Department of Education**; remove the federal government entirely
    - Put **all control in the hands of parents**, local school boards, and state government
    - **End all federal funding**

What is a centrist education policy?

- A broad mix of both left and right; **public/private partnership**

What is an ambiguous education policy?

- We need to "improve" education without a clear policy proposal
- Discussion of improving education, quality, but no discussion of government role

What is an irrelevant education policy?

- Discussions of school-to-prison pipeline, STEM, curriculum (beyond Common Core, NCLB)

Coding Energy Policy Statements (Renewables vs. Fossil Fuels)

This task involves reading text from congressional candidates' campaign platforms and judging how much they are left-leaning (very left, left) or right-leaning (very right, right). If you believe the text expresses a centrist position, select the "centrist" option. If you believe the paragraph does not express a clear position, select the "ambiguous" option. If you believe the text does not express a position relevant to immigration policy, select the "irrelevant" option.

What are "left" or "right" energy policies?

Very left positions tend to advocate for the following:

- **Full investment in renewable energy** (I am 100% committed…)
- Pass legislation to provide tax incentives, **ban fossils, subsidize green energy growth**
    - Rejoin Paris Climate Agreement; Green New Deal; net zero emissions by…

Left positions tend to advocate for the following:

- More in favor of renewable energy
- In favor of **increasing renewable energy,** committed to growing green energy
- Broad statements "I support" statements

Right positions tend to advocate for the following:

- More **in favor of fossil energy; market-driven energy policy**
- Protect the fossil energy sector; "we need fossil energy"
    - Reopen pipelines/fossil fuel facilities, open keystone pipeline
    - May discuss green energy as an unsuitable alternative (capacity, unproven, expensive)

Very right positions tend to advocate for the following (fossil, free market most important)

- **Fully investment in fossil fuels** (I am 100% committed…)
- **Remove regulations, abolish EPA, reduce government regulations on fossil fuels**

What is a centrist energy policy?

- "**All of the above approach**" to energy policy
- Discusses both fossil and renewable energy equally
- Says we must "balance" environmental and economic considerations
- Note: natural gas *is not* renewable

What is an ambiguous energy policy?

- Discusses mechanisms of technology (e.g., nuclear energy is safe, natural gas is clean…)
- We need more energy production; we need to invest in energy production

What is an irrelevant energy policy?

- Discussions of climate change *only*; environmental conservation *only*

Coding Gun Policy Statements (Restrition vs. Access)

This task involves reading text from congressional candidates' campaign platforms and judging how much they are left-leaning (very left, somewhat left) or right-leaning (very right, somewhat right). If you believe the text expresses a centrist position, select the "centrist" option. If you believe the paragraph does not express a clear position, select the "ambiguous" option. If you believe the text does not express a position relevant to immigration policy, select the "irrelevant" option.

What are "left" or "right" gun policies?

Very left positions tend to advocate for the following:

- **Purchase restrictions on guns** themselves (military-style, assault rifles, AK-47s, weapons of war)
- Repeal legal protections for **gun manufacturers, dealers**
- Ban all weapons

Left positions tend to advocate for the following:

- Some restrictions on **buying time; wait periods; raising the age of purchase**
    - When and where you can purchase (gun shows, wait periods, private sales)
- **Special permitting** for assault-style, certain weapons
- Creation of a national **registry** of gun owners
- Generally, make it harder to purchase guns, "stricter gun laws"

Right positions tend to advocate for the following:

- Discussion of support for 2nd amendment, but **no discussion of restrictions**
- Prevent liberals from taking guns away
- Against left gun positions (restrictions on where, when, and what can be bought)

Very right positions tend to advocate for the following:

- Repealing limitations on concealed carry, other gun access laws
    - **Reciprocity** in concealed carry across state lines
- Guns in schools, and college campuses (i.e., **"gun-free" zones**), provide teachers with guns
- No background checks, no waiting periods, no classes/permitting
- **Against centrist positions (basic, common-sense reforms)**
- No limits on guns whatsoever; **ANY gun control laws are a problem**

What is a centrist gun policy?

- **No-fly lists; red flag laws; mental health; closing loopholes; universal background checks**
- "Common sense" gun reforms
- **Ban on bump stocks, high-capacity magazines**
- Buy-back programs, with the option of keeping guns
- Can be pro 2nd Amendment, but mentions restrictions

What is an ambiguous gun policy?

- Discussions of responsible gun ownership; common-sense measures
- "Get guns off our streets", and "reduce gun violence" with no policy discussion

What is an irrelevant gun policy?

- Discussions of the prison system, cash bail

Coding Health Insurance Marketplace Statements (Public vs. Private)

This task involves reading text from congressional candidates' campaign platforms and judging how much they are left-leaning (very left, left) or right-leaning (very right, right). If you believe the text expresses a centrist position, select the "centrist" option. If you believe the text does not express a clear position, select the "ambiguous" option. If you believe the text does not express a position relevant to healthcare policy, select the "irrelevant" option.

What are "left" or "right" healthcare policies?

Very left positions tend to advocate for the following:

- **Full government involvement** in the health insurance market
    - **Single-payer, universal healthcare system, "Medicare-for-all"; abolish private insurance**
- Healthcare is a fundamental human right

Left positions tend to advocate for the following:

- **Expanded government involvement** in the health insurance market:
    - **Expand public options**; **expand Medicaid; expand Medicare**
    - May advocate for keeping private insurance for those who want it

Right positions tend to advocate for the following:

- **Little government involvement** in the health insurance market:
    - **Ensure preexisting condition coverage, cost caps on insurance**
- Advocates for choice on coverage but **does not go as far as to say no government involvement**
- **Free market**, purchasing across state lines, competition

Very right positions tend to advocate for the following:

- Remove government from health insurance marketplace, full free market
- **Repeal Obamacare; ACA; replace with competitive, free-market system**

What is a centrist healthcare policy?

- Maintain status quo; **protect** ACA; **protect** Medicare; **protect** Medicaid

What is an ambiguous healthcare policy?

- Broadly discusses reform to achieve affordability, lower cost, and greater access but **takes no explicit position**

What is an irrelevant healthcare policy?

- Discusses ONLY quality of healthcare, scope of services for individual populations, group-based
    - Black maternal mortality, veteran's healthcare, women's healthcare
    - Drug costs

Coding Immigration Policy Statements (Inclusive vs. Exclusive)

This task involves reading text from congressional candidates' campaign platforms and judging how much they are left-leaning (very left, somewhat left) or right-leaning (very right, somewhat right). If you believe the text expresses a centrist position, select the "centrist" option. If you believe the paragraph does not express a clear position, select the "ambiguous" option. If you believe the text does not express a position relevant to immigration policy, select the "irrelevant" option.

What are "left" or "right" immigration policies?

Very left positions tend to advocate for the following:

- Treatment of immigrants: **pathway to citizenship for *all* undocumented**
- Calls to end all deportation
    - Provide healthcare, education, humane treatment
- Border Control, Security: **Abolish, demilitarize ICE**
- General Approach: inclusive; maximize access to immigration

Left positions tend to advocate for the following:

- Treatment of immigrants**: some pathway** to citizenship (e.g., **DACA only**)
    - Earned pathway: military service, civil service; provide temporary protected status (TPS)
- Border Control, Security: **No kids in cages, humane treatment** at border
- General Approach: inclusive; expand access to immigration
    - A mix of law + citizenship, with **more emphasis on pathways to citizenship**

Right positions tend to advocate for the following:

- Treatment of immigrants: **emphasize lawful immigration**; illegal immigrants must go through the regular immigration process, with no special treatment
- Recognize the need for seasonal, temporary workers; keeping "talented" illegal immigrants
- Border Control: **Secure the border;** investment in security
- General Approach: exclusive; reduce access to immigration
    - A mix of law + citizenship, with **more emphasis on security**

Very right positions tend to advocate for the following:

- Treatment of immigrants: **all illegal immigrants as criminals; no amnesty; deport all illegal**
- increased limits on who can immigrate (support for Muslim ban, only high-skilled workers, country-specific restrictions); **end birthright citizenship; end chain migration**
- Border Control: Heavy investment in border security
- **Close borders, end all immigration**

What is a centrist immigration policy?

- An **even mix of enforcing laws at the border + some pathway to citizenship**

What is an ambiguous immigration policy?

- Just platitudes: we are a nation of immigrants, treat everyone with respect

What is an irrelevant immigration policy?

- None