Measuring Strategic Positioning in Congressional Elections

Colin R. Case[*]

June 1, 2023

Abstract

Measuring strategic ideological positioning is fundamental to the study of representation in political science. But current measures of candidate positioning present limitations for the study of candidate behavior and polarization in the American context. I introduce Website EmBedding (WEB) Strategic Positioning Scores. WEB Scores employ word embeddings with document-level vectors trained on congressional candidates' issue statements, as presented on their campaign websites. The estimates produced from this procedure have high construct validity and improve upon current measurement limitations, including expanding the number of candidates with estimates and using actual issue-position data to produce estimates. Further, the embeddings for each candidate present a rich representation of underlying issue text and offer a variety of avenues for future research. Demonstrating the measurement's utility, I show that incumbent candidates respond to the positioning of primary challengers by changing their own positioning. I conclude by providing recommendations for choosing a measure of candidate positioning.

[*]The University of North Carolina at Chapel Hill, Department of Political Science. I am grateful for helpful comments and feedback from Max Goplerud, Eric Hansen, Jacob Montgomery, Emily Ommundsen, Rachel Porter, Jason Roberts, Brandon Stewart, and Sarah Treul. I would also like to thank conference participants at PolMeth 2022 and the Junior Americanist Workshop Series MPSA Mini-Conference for their engagement and feedback.

In 2020, Representative Andy Kim ran unopposed in the Democratic primary for New Jersey's 3rd congressional district. While Kim advocated for policies consistent with the Democratic Party, his proposals were not extreme by any means; Kim suggested "pragmatic" solutions to address climate change, "bipartisan" proposals to rebuild America's infrastructure, and "building" on the Affordable Care Act. In 2022, however, Rep. Kim was challenged in the Democratic primary by Reuven Hendler, a first-time candidate running for office. Hendler ran far to the left of Kim, strongly advocating for a Medicare-for-all healthcare system, tuition-free college, and aggressive policies to combat climate change. Downsian theories of candidate placement suggest Kim should consider adopting more liberal policies in the 2022 election. With existing measures, however, it is not possible to test whether Hendler's positioning resulted in Rep. Kim changing his position–Hendler did not receive any donations, he has no prior political experience, and he did not win the primary election.

Research on congressional elections is often interested in the effects of strategic ideological positioning, such as whether candidates who position themselves at the extreme are more successful in elections (e.g., Hall 2015; Hall and Snyder 2015). However, current measurement approaches (e.g., Poole and Rosenthal 1985; Clinton, Jackman and Rivers 2004; Bonica 2014; Christopher et al. 2015; Barberá 2015; Macdonald et al. 2022; Gaynor et al. 2022) suffer from two primary limitations. First, large populations of candidates with substantive importance (e.g., primary candidates, candidates without elected experience, hopeless candidates), such as Reuven Hendler, are often excluded from existing measurements. Second, existing approaches capture related, but conceptually distinct, quantities of interest. For example, a number of measures rely on voters' perceptions of candidates' positions, not the actual issues candidates run on. Both limitations place a significant restriction on the types of research questions that can be asked about candidate positioning in congressional elections.

Given the shortcomings described above, I propose a new measure of congressional candidate positioning, Website EmBedding (WEB) Strategic Positioning Scores, using data collected from campaign website issue pages by Porter, Treul and Case (2023) for the 2018, 2020, and 2022 primaries for the U.S. House of Representatives. Campaign websites are well-situated to study primary candidate positioning: they are unmediated, in that they come directly from the campaign; not subject to other gatekeeping (e.g., media); and contain a range of policy areas (Druckman, Kifer and Parkin 2009). Moreover, campaign websites mitigate the limitations presented above by providing better coverage than existing measures and are based on the actual positions candidates take during the election.

To estimate WEB Scores, I rely on recent developments in word embedding models that allow for the inclusion of a document-level vector for each candidate-year occurrence. Rheault and Cochrane (2020) validate this approach across various contexts as a suitable way to uncover elite positioning. Because this model captures word embeddings and candidate embeddings in the same dimensional space, it is possible to carry out numerous validation procedures between the measurement and underlying text. I also demonstrate the measure's utility in answering new questions given the expanded number of candidates with a position estimate. Using this measure, I show incumbent candidates challenged by an extreme (moderate) challenger position become more extreme (moderate) in their ideological positioning.

The paper proceeds as follows: I first outline existing measures of strategic candidate positioning and their limitations. Turning to campaign websites, I demonstrate the extent to which data from campaign websites improve upon existing measures. I discuss the estimation strategy and conduct robustness checks to ensure WEB Scores are not sensitive to modeling decisions. I then validate the resulting measure against existing measures of candidate positioning and show WEB Scores have high construct validity. I also take advantage of the properties of word embedding models and show that WEB Scores capture semantic relationships in text that reflect various positions candidates can take. I conclude by demonstrating WEB Scores' utility and providing recommendations to researchers using these measures in substantive research on candidate positioning and congressional elections.

Measuring Strategic Ideological Positioning

Measuring candidates' strategic ideological positioning is crucial to studying representation and elections in political science. Much of this work is grounded in the theory that voters prefer candidates proximal to their position (Black 1948; Downs 1957). In this manner, candidates have an incentive to take into consideration the positions of both their competitors and the electorate when deciding where to position themselves on the issues. Strategic ideological positioning can, therefore, be defined as the aggregate set of positions candidates take, that are constrained not only by a latent set of preferences, but by other electoral dynamics that candidates respond to in order to maximize their electoral chances. Research has attempted to measure the placement of members based on their roll-call votes, with the most widely used measure being DW-NOMINATE (Poole and Rosenthal 1985). Other approaches have also used roll-call voting measures with different methodological approaches (Clinton, Jackman and Rivers 2004) or assumptions (Duck-Mayr and Montgomery 2022).

While these approaches provide a theoretically and empirically important perspective, measures based on roll-call votes present limitations for studying candidate positioning in congressional elections. Namely these measures only focus on legislative position taking for select issues that make it to the floor, not the universe of issues candidates run on. Moreover, as this behavior is only observed for members of Congress, these existing measures make it impossible to evaluate the positions of non-incumbent candidates. As an alternative, recent work has developed a number of measurement strategies to capture candidate positioning in congressional elections. Broadly speaking, these election-focused measurement strategies approximate positioning relying on one of two sources: either citizen perceptions of candidates or actual candidate behavior.

Within the first typology, measures based on citizen perceptions rely on the assumption that citizens are able to take into account candidate positioning, such as the issues they run on, policy goals, and values (Bonica 2014). Common approaches often ask survey respondents (Christopher et al. 2015; Ramey 2016) or experts (Hirano et al. 2015) to place candidates spatially from liberal to conservative and aggregate these responses to position candidates using various scaling methods. Other approaches rely on aggregate citizen behavior, such as donations (Bonica 2014) or follows on Twitter (Barberá 2015). These estimation strategies assume that citizens donate to and follow candidates on Twitter who are positioned proximally to one another.

The second typology of measurement approaches approximates candidate positioning using other related candidate behaviors. For example, Macdonald et al. (2022) use news story domain sharing (e.g., Fox News or CNN) on Twitter for members of Congress to spatially place candidates. Other approaches, such as that used by Gaynor et al. (2022), employ text-based scaling of members of Congress across a variety of different contexts, including tweets and floor speeches. Another subset of measurement strategies focuses specifically on state legislators. For example, Ansolabehere, Snyder and Stewart (2001) and Montagnes and Rogowski (2015) use Project Vote Smart's NPAT survey of state legislators while Shor and McCarty (2011) rely on roll-call votes in state legislative bodies and use the NPAT survey to link state legislators across institutional contexts.

Across both measurement typologies, two issues persist. First, these measurement approaches exclude a large proportion of candidates (e.g., primary candidates) who are substantively relevant to a variety of research questions. Second, these measurement strategies rarely capture the actual issues candidates are running on, posing problems for inference and measurement validity. In the following subsections, I outline the limitations of the various measurements discussed above as well as the constraints placed on the types of substantive questions that can be asked.

Measurement Coverage

The vast majority of measurement approaches exclude a large proportion of candidates–oftentimes those with a lower chance of winning elections or those who do not have previous political experience. Across different measures, the excluded population differs. Measures that rely on survey responses from voters are often limited to general-election candidates. This is due, in part, to resource constraints: asking about 2,000+ candidates who run in congressional primaries is not feasible. There are also knowledge limitations as well; it is unlikely the average voter is aware of the positions of all candidates running in a primary race (see Ahler, Citrin and Lenz 2016). Similarly, measures employing experts to place candidates from liberal to conservative often focus on general-election candidates or certain high-profile races due to the same resource and knowledge constraints. These constraints place limitations on the study of positioning in congressional primaries. In the last decade, however, the vast majority of congressional elections are considered "safe" districts, with little chance of the opposing party posing a threat in the general election (Jacobson and Carson 2016). In these districts, electoral competition is centered in the primary. Estimating the positioning of candidates in primary elections is, therefore, crucial for a variety of research questions related to candidate emergence and success (e.g., Thomsen 2014; Hall and Snyder 2015; Thomsen and Hall 2018).

Other measurement approaches, such as CFscores (Bonica 2014) or state legislator scores (Shor and McCarty 2011), do include a subset of primary candidates, but still exclude substantively important populations of candidates. Most often, the excluded groups of candidates are those without political experience (in the case of state legislator focused measures) or those with little chance of winning the election (in the case of donation-based measures).[1] Both subsets of excluded primary candidates are relevant to understanding election outcomes, candidate emergence, and the responsiveness of incumbents to primary challengers.

Limitations in candidate coverage restrict the types of research questions that can be asked. For example, Treul and Hansen (2023) evaluate the extent to which working-class candidates who run can win primary elections. After controlling for relevant factors, they find that working-class candidates are less likely to win and receive a lower vote share. In the paper, the authors cannot control for candidate ideology, a variable commonly used in a model for predicting candidate success. Due to the fact that 63% of working-class candidates do not receive enough donations to have a CFscore. Research designs, like Treul

[1]For context, CFscores, which provide some of the highest levels of coverage of primary candidates among existing measures, did not have a score for 1,386 (35.5%) candidates for the U.S. House of Representatives in 2018 and 2020.

and Hansen's, where certain subsets of candidates are the focus, are often not able to consider the positioning of these candidates to better understand their lack of success in primary elections.

In addition to the above research limitations, the lack of coverage among inexperienced candidates limits the scholarly understanding of incumbent positioning. Members of Congress are responsive to the challengers they face and change their positions as such (Sulkin 2005). This plays out both in the legislative aspect of a member's job (Jewitt and Treul 2019) and the issues candidates run on (Porter, McDonald and Treul 2021). But as existing measures of positioning exclude a large subset of candidates, researchers are not able to consider the effects of these excluded lesser-known challengers running against incumbent candidates.

Underlying Data

Beyond coverage challenges, previous measures often rely on approximations of candidate positing, whether that be through perceptions that citizens form (e.g., Bonica 2014; Christopher et al. 2015; Hirano et al. 2015; Ramey 2016) or other behaviors associated with the positions candidates take (e.g., Macdonald et al. 2022; Montagnes and Rogowski 2015; Shor and McCarty 2011). While both types of measurement are theoretically related to candidate positioning, there is reason to suspect these measurements may be capturing distinct concepts. For example, when it comes to citizens' perceptions, other factors can cause perceptions and actual positioning to diverge. As Bonica (2014) notes, one example of this are endorsements that can sway donors' perceptions of candidates. But endorsements can happen for reasons unrelated to a candidates' positioning, such as likelihood of winning the election. While these measurement strategies do capture fundamentally important concepts—such as where voters' perceive candidates to be positioned—it is theoretically related but conceptually distinct measure from candidates actual issue positioning.

By relying on approximations, previous measurement strategies also make it difficult to connect the measurement to changes in behavior. For example, in recent congressional elections, the intra-party correlation between CFscores, a widely used measure of positioning based on who donates to campaigns, and DW-NOMINATE diverges significantly (see Barber 2022). While Barber (2022) investigates several potential reasons for this divergence, such as institutional, regional, and racial differences, as well as changes in donor behavior, the data-generating process underlying many approximations for candidate positioning make it difficult to adjudicate the reason for measurement divergence, as there is no underlying candidate behavior that can be connected back to validate changes in the measurement.

Furthermore, by not capturing actual candidate positions, previous measures, specifically those based on citizen perceptions, can pose threats to inference for certain research questions. Oftentimes, scholars are interested in how candidate positioning is related to other candidate behaviors. Consider the following directed acyclic graph where a candidate's position is not directly captured, but related to the subsequent measure of that position, as is the case with the measures discussed above. In this general theoretical framework, the research question focuses on the relationship between the measure of candidate positioning and other candidate behaviors. If it is the case that the subsequent candidate behavior also shapes citizens' perceptions, a measure of candidate positioning based on citizen perceptions becomes endogenous.

$$\begin{tikzcd} \text{Candidate Positioning} \arrow[r] & \text{Measure of Positioning} \arrow[r] & \text{Candidate Behavior} \arrow[r] & \text{Citizen Perceptions} \arrow[ll, bend right, dashed] \end{tikzcd}$$

For example, Case (2023) argues that candidates use the visual elements of their campaign, such as the colors in their logo, to convey information to voters about their positioning. Subsequently, these visual elements of a campaign affect voters' perceptions of candidates' positioning. In this instance, a measure based on perceptions of candidates would be endogenous because the link between the measure of candidate positioning and candidate behavior would be an artifact of the measurement strategy. This again places limitations on the types of research questions that can be asked with existing measures.

Data Description

To improve upon the limitations of current measurement approaches, I propose campaign website issue positions as an alternative data source for estimating candidate positioning. Websites are an important part of the modern candidate's campaign. Most candidates in recent years (88.1% between 2018-2022) maintain a website that acts as an "information hub" for all parts of the campaign, from information about the candidate to their issue positions and policy proposals (Herrnson, Panagopoulos and Bailey 2019). Candidates carefully craft these websites, knowing that potential voters, donors, journalists, and other electoral stakeholders will visit them for information about their campaign (Druckman, Kifer and Parkin 2009). These websites come directly from their campaign, cover a range of issues and policy areas, and are representative of the population of campaigns (Druckman, Kifer and Parkin 2009). Further, throughout an election cycle, little changes on the campaign website, making it a static representation of a campaign for that election (Porter, McDonald and Treul 2021). To this extent, campaign websites are a comprehensive data source for studying candidates in U.S. congressional elections.

As part of their campaign website, most (76.6%) candidates maintain an "issue page" that explicitly lays out the candidate's stance on the issues, specific policy proposals, and oftentimes commentary on contemporary events. Porter, Treul and Case (2023) collect the issue pages for all Democratic and Republican primary candidates for U.S. House of Representatives who had an official campaign website in 2018, 2020, and 2022. To collect official campaign website issue positions data, Porter, Treul and Case (2023) first identified the names of all candidates running in the primaries from state election boards as well as candidate filings with the Federal Election Commission (FEC). Using this list of names, official campaign websites were identified using a few different sources. `Politics1.com` maintains a database of all campaign websites for candidates running actively in each race and is where the links to a majority of the campaign websites were found. Others were found through various social media pages, `Ballotpedia.com`, and Google searches.

As a part of this data-collection process, research assistants identified whether or not each candidate had a "platform," or a set of issue statements.[2] While this looks different on some websites, it oftentimes is referred to as "My Platform," "Issues," or "Where I Stand." On these issue pages, candidates typically organize their issue stances in a series of paragraphs about different policy areas, or individual issue statements. Research assistants manually collected each of these individual issue statements. This process was done contemporaneously in the ten days leading up to each candidate's primary both to ensure consistency in candidates in the data collection process and that candidates websites were finalized in the lead up to the election. In total, this data set contains 4,554 issue pages (76.6% of all candidates; 86.9% of candidates with a website). Across 2018, 2020, and 2022, each candidate had a mean of 9.8 individual issue statements on their campaign platform for a total of 44,544 individual issue statements. For a full discussion of the data and the data collection process, see Porter, McDonald and Treul (2021).[3]

Campaign issue pages improve upon current measurement approaches through both the expansion of the number of candidates included and through a closer connection of the underlying data to the theoretical quantity of interest. To compare the coverage of candidates with an issue page versus previous measurement approaches, Figure 1 plots the number of candidates with an issue page versus the number of candidates with a CFscore for the 2018 and 2020 U.S. House of Representatives primaries.[4] This is further broken down by candidate

[2]Screen shots of example campaign platforms can be found in Appendix A.

[3]Interviews with campaign consultants who work with candidates on setting up their website highlighted the importance of these pages, mentioning issue pages as the part of the campaign they spent the most time discussing with candidates. It should be noted that while these campaign consultants often use similar strategies across campaigns (Nyhan and Montgomery 2015), interviews highlighted a few important components that ensure the website is capturing candidate behavior. First, while campaign consultants help with the drafting process of issue pages, it is still what the candidate is interested in and wants to focus on for the election that shapes the issue pages. Second, candidates are still the ones operating their campaign, and even with the direction of campaign consultants, the candidate is the one with the final say. Third, despite consistent strategies across the same consulting firms, most have a review process across candidates to ensure that issue text for one candidate is not the same as issue text from another candidate at the same firm; most of the time this process involves separate writers for the issue pages and a secondary check of all issue text. In this manner, these issue pages are individual to each candidate.

[4]2018 and 2020 are used for Figure 1 because these are the current years for which CFscores are available.

type: incumbents, those who have previously held elected office, and those who have not previously held elected office. In the aggregate, 2,941 (75.4%) candidates had an issue page on their campaign website in 2018 and 2020 and 2,514 (64.5%) have a CFscore.[5] As is evident in Figure 1, campaign websites provide a large increase in coverage of candidates when it comes to inexperienced challengers. Of the 2,654 inexperienced candidates who ran in 2018 and 2020, 1,856 (70.0%) had an issue page on their campaign website, while only 1,366 (51.5%) received enough eligible contributions for a CFscore. When it comes to experienced challengers, both sets of data have a high percentage of candidates, with 367 (76.0%) having an issue page and 392 (81.1%) out of 483 total experienced challengers having a CFscore in 2018 and 2020. Importantly, it should be noted that a small number of incumbents do not have an issue page on their campaign website, leading to slightly worse coverage with campaign websites (94%) than CFscores (100%).[6]

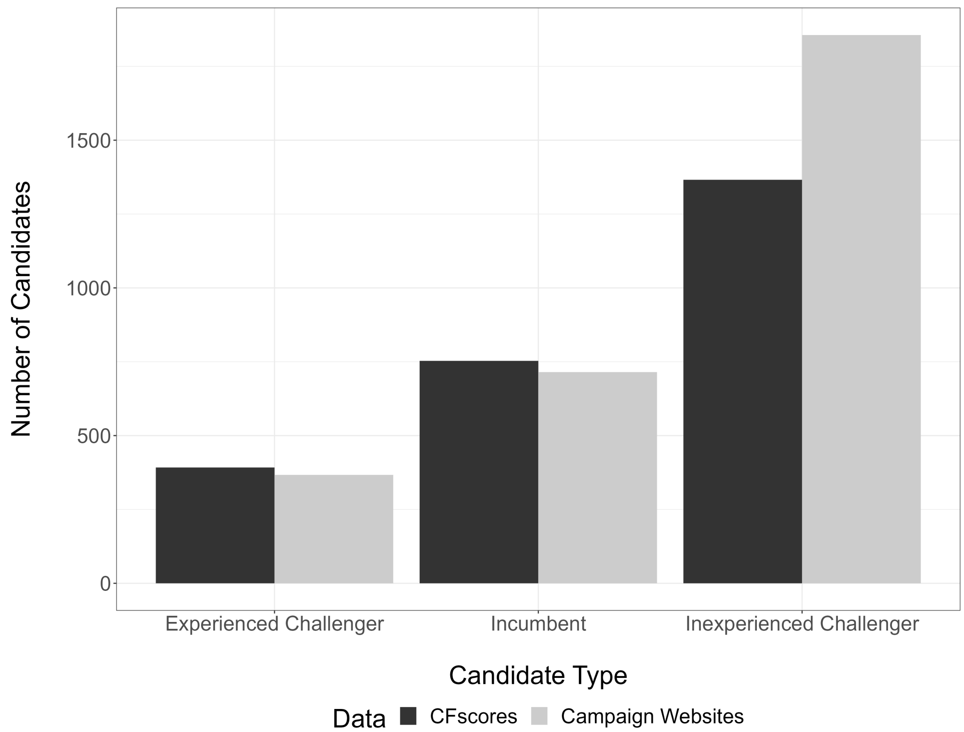
Moreover, campaign websites do well capturing candidates who did not receive enough eligible contributions for a CFscore. In 2018 and 2020, across all categories of candidates, there were 1,386 candidates without a CFscore. Among those, 745 (53.8%) have a campaign website issue page. For research that is substantively interested in candidates who may not receive many donations, such as Treul and Hansen (2023), or work interested in how primary challengers can shape member behavior (e.g., Jewitt and Treul 2019; Porter, McDonald and Treul 2021), using campaign websites as a data source represents a substantial improvement in the ability to test hypotheses related to candidate positioning. Using underlying data that increases the number of candidates with a valid position estimate is an important advantage of campaign websites over other measures of strategic ideological positioning.

Patterns for campaign websites are consistent in 2022.

[5] CFscores are used as a comparison measurement due to the high level of coverage compared with other measures of candidate positioning. Other measures of candidate positioning have a substantially lower percentage of candidates included.

[6] While the number of incumbents without an issue page is a limitation of the measurement, it should be noted that the number of incumbents with websites and issue pages has increased in the most recent election cycles. This is also true for across other candidate backgrounds. In 2022, coverage of candidates is 366 (97.3%) for incumbents, 225 (79.5%) for experienced challengers, and 1,022 (73.7%) for inexperienced challengers.

Figure 1: Candidate Coverage by Measurement and Candidate Type



Note: Figure 1 depicts the number of candidates running as either a Democrat or Republican in 2018 and 2020 congressional primary elections who have a CFscore (black bar) and an issue page on their campaign website (gray bar). The data are broken down by challengers with previous elected experience (left), incumbent candidates (middle), and challengers with no previous elected experience.

Estimation Strategy

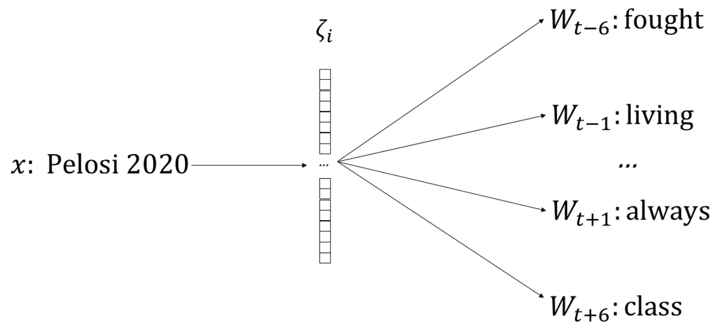
To estimate Website EmBedding (WEB) Strategic Positioning Scores using issue position text, I rely on word embedding models. Word embeddings are the parameter estimates from neural network models designed to predict word(s) given the context around that word(s). Work in computer science has highlighted the different ways in which word embeddings can capture important underlying properties of language, such as the similarity between words, analogies, and antonyms (Mikolov, Yih and Zweig 2013). Word embedding models have recently seen more wide-spread use in a political science context (Rodriguez and Spirling 2022). Their rise in use stems from the ability to assess and test hypotheses for how word use can differ across covariates (Rodriguez, Spirling and Stewart 2021) as well as uncover important latent traits related to the properties of both words (Grand et al. 2022) and the people using them (Rheault and Cochrane 2020). Moreover, Rodriguez and Spirling

(2022) show that word embedding models are able to identify nearest neighbors to politically relevant terms, such as immigration, at the same level as human coders. This suggests embeddings are well-suited to pick up on important semantic relationships in text related to political phenomenon such as candidate positioning.

There are a number of other approaches for measuring candidate positioning in text. Among the earliest approaches to estimate positioning using text are WordScores (Laver, Benoit and Garry 2003). WordScores are a supervised machine-learning method that uses a smaller sample of labeled documents, where each document has been labeled by experts to identify their positional leaning. Based on the occurrence of each word in the labeled documents, words then receive a score representing their positional lean. From there, unlabeled documents can then receive a placement based on the occurrence of words and the scores for each word from the previous. WordFish (Slapin and Proksch 2008) is another method that similarly relies on the occurrence of words in a document. Instead of using labeled documents, WordFish uses regressions to project counts for each word onto each party-year combination. More recently, Vafa, Naidu and Blei (2020) develop text-based ideal points (TBIP) that also uncover specific topics associated with each latent score, providing more validity and taking into account the co-occurrence of words.

While WordFish and TBIP improve upon supervised methods by reducing the time and cost of labeling documents, all three methods still rely on the occurrence (or co-occurrence in the case of TBIP) of words in a document without taking into account the full context of word usage. This contributes to these models having little sense about the semantic relationship between words (e.g., "powerful," "strong," and "Paris" are equally distant semantically) after the model is estimated (Le and Mikolov 2014). This is an important point when estimating candidate positioning from text. For example, take the words "immigration" and "immigrants." While different parts of speech, both words are semantically similar. An estimation strategy strictly relying on the occurrence of words is not able to account for the semantic similarity between these words. Word embedding models improve upon this limitation in their ability to incorporate high quality semantic relationships between words

Figure 2: Model Architecture with Window of 6: Candidate-Year Embedding Estimation

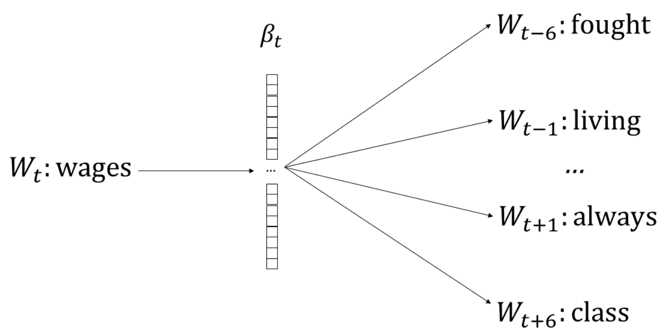


during the training process.

Model Architecture

To estimate WEB Scores, I rely on a word embedding model with document-level vectors (Doc2Vec) for each candidate-year. Prior work has shown the utility of document-level vectors in capturing important latent political constructs, such as positioning (Rheault and Cochrane 2020). In the model, each candidate-year, i, and word, j, has an embedding with M dimensions, denoted as ζ_i and β_j respectively. The model has two parts. The first part is the training process for each candidate-year embedding and follows the Paragraph Vector Distributed Bag of Words (PV-DBOW) approach developed by Le and Mikolov (2014). In this estimation, for each document, a word, w_t is sampled at each iteration and the window, Δ, before and after that word is extracted. The resulting window surrounding w_t, denoted as w_Δ, are the outcomes of interest. The output can be written more completely as $w_\Delta = (w_{t-\Delta}, \ldots, w_{t-1}, w_{t+1}, \ldots, w_{t+\Delta})$. The model input is an indicator vector, x_i for the candidate-year author of the document. x_i is multiplied by the matrix of candidate embeddings, ζ. The resulting candidate-year embedding, ζ_i is used to predict each word in the window using a softmax classification between ζ_i and β_j. The parameters for the embeddings are then fitted by minimizing the cross-entropy loss using stochastic gradient descent. A graphical depiction of this process is in Figure 2.

Figure 3: Model Architecture with Window of 6: Word Embedding Estimation



It is important to note in this part of the model, w_t is not used to predict the window like it would be in a traditional skip-gram word embedding model–only the candidate-year embedding is. Further, in the traditional PV-DBOW approach, word embeddings are not stored. While this leads to a more efficient estimation (see Le and Mikolov 2014), the quality of the results is often inconsistent (Lau and Baldwin 2016). For this reason, I follow Lau and Baldwin (2016) and use a simultaneous skip-gram word embedding model that trains word embeddings in the estimation procedure. This represents the second part of the model. The model architecture is the same as Figure 2, but instead uses the target word, w_t, as the model input. The indicator vector for w_t is multiplied by the matrix of word embeddings and the resulting word embedding for w_t, β_t, is used to predict each word in the window through softmax classification (see Figure 3). With this simultaneous estimation, the second part of the model is indirectly affecting the candidate-year embedding training by developing word embeddings that capture the semantic relationship between politically relevant terms. This simultaneous estimation significantly outperforms traditional document vector models as well as other N-gram models (Lau and Baldwin 2016). Further, the use of a skip-gram word embedding model leads to better performance in building semantic pairings, rather than syntactic pairings (Mikolov, Yih and Zweig 2013) that are more important for capturing candidate positioning in text.

Model Implementation and Fit

Before fitting the word embedding model on campaign website text, it is important to discuss a number of parameter-level decisions in creating the resulting WEB Scores. Starting with text pre-processing, I follow the same procedure as Rodriguez and Spirling (2022) and convert all tokens to lower case and remove all non-text characters. In addition, I also remove words that do not appear across the full set of documents more than five times. This is done because Doc2Vec uses an estimation strategy that generally over-weights rare terms in the training process.[7] Removing infrequent terms improves the accuracy and performance of the models (Rodriguez and Spirling 2022).

I fit the model using parameter recommendations from Rodriguez and Spirling (2022), including a window of 6 and an embedding dimension of 300.[8] In addition, I also use pre-trained Word2Vec embeddings from the Google News corpus. This is done due to the limited data from campaign issue statements for training a word embedding model and ensures high-quality word embeddings are used in the training process. The use of pre-trained embeddings also improves the performance of Doc2Vec embedding models Lau and Baldwin (2016). Finally, I use default hyperparameter recommendations from Mikolov, Yih and Zweig (2013) with an increased number of epochs (20). The larger number of epochs is due to a limited number of documents per candidate-year occurrence and is consistent with Rheault and Cochrane (2020).

From the model output, each word and candidate-year has an embedding of 300 dimensions. While these embeddings represent a rich understanding of the syntactic and semantic relationship between words and candidates, applied researchers are often interested in univariate measures of positioning that can be used in regression analysis. To produce the resulting WEB Scores, I follow Rheault and Cochrane (2020) and use principal component analysis to reduce the candidate embeddings. In determining the number of dimensions, I

[7]Estimates of candidate positioning are highly correlated with different cutoff thresholds (0, 5, 10, 20).

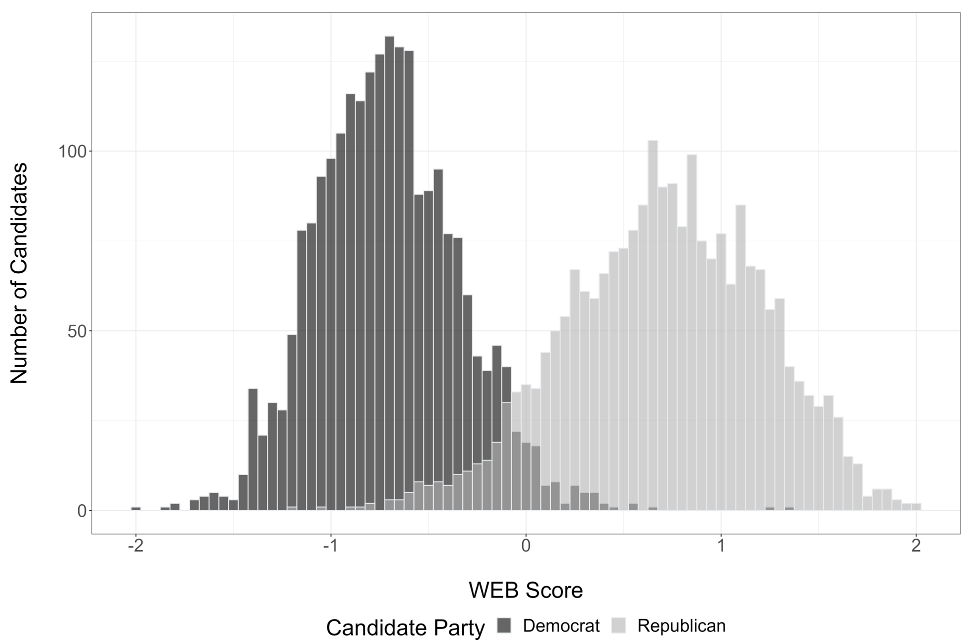
[8]Because there is no clear-cut justification for model parameters, I also fit models with various window sizes (5, 6, 7, and 8) and embedding dimensions (100, 200, 300) and show that resulting measures are almost perfectly correlated (≥ 0.99), suggesting parameter decisions have little effect on the resulting scores. See Appendix B for correlation tables from different model architectures.

identify the knee point using the Kneedle algorithm (Satopaa et al. 2011). The algorithm identifies a single dimension as the inflection point resulting in a WEB Score for each candidate in each election.

Analysis

In this section, I provide an overview of the measurement and several validation procedures. The distribution of WEB Scores is plotted in Figure 4. The measurement has a mean of 0 across candidates and a standard deviation of 0.83. Democratic candidates trend to the negative side of the scale with a mean score of -0.71 while Republicans have a mean score of 0.71.[9]

Figure 4: Histogram of WEB Scores in 2018, 2020, and 2022 U.S. House of Representatives Primary Elections



Note: Figure shows the distribution of candidates by WEB Score. Democratic candidates are colored dark gray and Republican candidates are colored light gray. Negative values represent more liberal positioning scores and positive values represent more conservative positioning scores.

To provide context to the measurement, Table 1 presents the ten most liberal and conservative incumbent candidates for the U.S. House of Representatives from 2018-2022. Notable

[9] From the model output, the axis is flipped to follow standard convention with Democrats on the left and Republicans on the right.

Table 1: Most Liberal and Conservative Incumbent Candidates

Most Liberal	Most Conservative
Barbera Lee (2020)	Troy Bladerson (2020)
Barbera Lee (2022)	Jody Hice (2020)
Nydia Velazquez (2022)	Neal Dunn (2020)
Yvette Clarke (2022)	Neal Dunn (2022)
Nydia Velazquez (2020)	Jeff Duncan (2020)
Melanie Stansbury (2022)	Neal Dunn (2018)
Suzanne Bonamici (2022)	Greg Murphy (2020)
Marilyn Strickland (2020)	William Timmons (2022)
Earl Blumenauer (2020)	Troy Balderson (2018)
Kathy Manning (2022)	Jeff Duncan (2022)

candidates such as Alexandria Ocasio-Cortez (2018) sit well to the left of the Democratic mean, with a score of -1.31. On the Republican side, Marjorie Taylor Greene (2020) also has a score well to the extreme of the Republican mean at 1.66.

In the sections to follow, I carry out a variety of tests to evaluate the validity of the measurement. I first assess the external validity by comparing the measurement with other measures capturing facets of positioning (DW-NOMINATE and CFscores) and show the measures are highly correlated. I also show the measure picks up on various intra-party differences by comparing scoring across various ideological caucuses, a tool members of Congress use to convey their position within parties to voters (Clarke 2020). In addition, I take advantage of the properties of embeddings to show that candidate positioning scores pick up important semantic differences that are related to the positions candidates can take on issues and are reflective of the scale endpoints.

External Validity

To assess the external validity of the WEB Scores, I first evaluate similarity between WEB Scores and pre-existing scores of positioning: CFscores, which capture donors' perceptions of candidates' positioning, and DW-NOMINATE, which captures voting preferences on the congressional legislative agenda. It should be noted that while these concepts are distinct from explicit candidate positioning, they should nonetheless be related. I focus these validation

Table 2: WEB Score and CFscore Correlations

	Correlation
All Candidates	0.86
Democrats	0.24
Republicans	0.21

tests on 2018 and 2020 candidates, the election years with both CFscores and DW-Nominate. The correlations are broken down into two tables. Table 2 shows the correlations (both all candidates and intra-party) for candidates with both a CFscore and a WEB Score. Table 3 reflects the same comparisons for candidates who were elected to Congress. I also include correlations between CFscores and DW-NOMINATE as a baseline for comparison over the same period with the same set of candidates.

Starting with Table 2, correlations with all candidates are high between CFscores and WEB Scores at 0.86. It should be noted, one of the differences between CFscores and WEB Scores is the amount of overlap between candidates from each party. According to CFscores, very few Democrats in 2018 and 2020 were more conservative than the mean; the same holds for Republicans and being more liberal than the mean. As is evident in Figure 4, WEB Scores have some cross-party overlap. This is consistent with other text-based scaling measures (e.g., Gaynor et al. 2022). In a polarized era, intra-party correlations are important to validating a score (Tausanovitch and Warshaw 2017). Intra-party correlations are weakly correlated at best, with a correlation of 0.24 for Democratic candidates and 0.21 for Republican candidates. In the aggregate, the results suggest that both CFscores and WEB Scores are capturing distinct concepts within parties, and further validation is needed.

Table 3 shows the correlations for candidates elected to Congress in 2018 and 2020, restricted to those having a WEB Score, a CFscore, and a DW-Nominate score in the 116th and 117th Congress. The first panel looks at all candidates, the second looks at Democratic candidates, and the third looks at Republican candidates. Starting with all candidates, the correlation between WEB Scores and DW-Nominate is high at 0.89, similar to the correlation between DW-Nominate and CFscores at 0.92. Turning to intra-party correlations for Demo-

Table 3: Measure Correlations for 116th and 117th Congress

All Members of Congress			
	CFscores	DW-Nominate	WEB Scores
CFscores	–	0.93	0.88
DW-Nominate	0.93	–	0.89
WEB Scores	0.88	0.89	–
Democrats			
	CFscores	DW-Nominate	WEB Scores
CFscores	–	0.02	0.43
DW-Nominate	0.02	–	0.21
WEB Scores	0.43	0.21	–
Republicans			
	CFscores	DW-Nominate	WEB Scores
CFscores	–	0.53	0.28
DW-Nominate	0.53	–	0.48
Web Scores	0.28	0.48	–

cratic candidates, WEB Scores are weakly correlated with DW-Nominate at 0.21.[10] This is higher than the correlation between CFscores and DW-Nominate for Democrats (0.02). It should also be noted that the correlation between WEB Scores and CFscores increase when only looking at candidates elected to Congress (0.43) compared with all candidates (0.24 from Table 2). Among Republican candidates who were elected to Congress, the correlation between WEB Scores and DW-Nominate is moderate at 0.48. This is, again, similar to the intra-party correlations for CFscores and DW-Nominate at 0.53. In the aggregate, all three measures are related when focusing on all members of Congress, and both diverge slightly when looking at intra-party correlations, suggesting they are capturing distinct, but related, constructs.

To further validate the measure against previously established proxies for candidate positioning, I also compare the average WEB Score for each ideological caucus in Congress. Members view ideological caucuses as a means to convey their positioning to donors and voters, especially within parties (Clarke 2020). Considering that joining an ideological caucus

[10]It should be noted that part of this can likely be attributed to shortcomings of DW-Nominate in correctly identifying more liberal members of Congress. Correlations with alternative estimation strategies, such as Duck-Mayr and Montgomery (2022), would likely be higher within the Democratic Party given this adjustment.

can be motivated by electoral interests, WEB scores should reflect positioning differences among caucus groups *within* parties.

To do this, I collect ideological caucus membership data for six caucuses that are relevant for the 2018, 2020, and 2022 elections (from most liberal to most conservative, according to Clarke (2020)): the Progressive Caucus, the New Democratic Coalition, the Blue Dog Coalition, the Main Street Partnership, the Republican Study Committee, and the Freedom Caucus. Caucus membership is collected for incumbent candidates during the Congress that runs concurrently with the election (e.g., incumbent candidates running in the 2018 election and caucus membership in the 115th Congress running from 2017 to 2019). Data for the 115th Congress comes from Clarke (2020) while the 116th and 117th Congresses were collected from archived official caucus membership pages. Because the Freedom Caucus does not maintain an official caucus roster online, the membership was gathered from a news article[11] that provided a roster of Freedom Caucus members.[12]

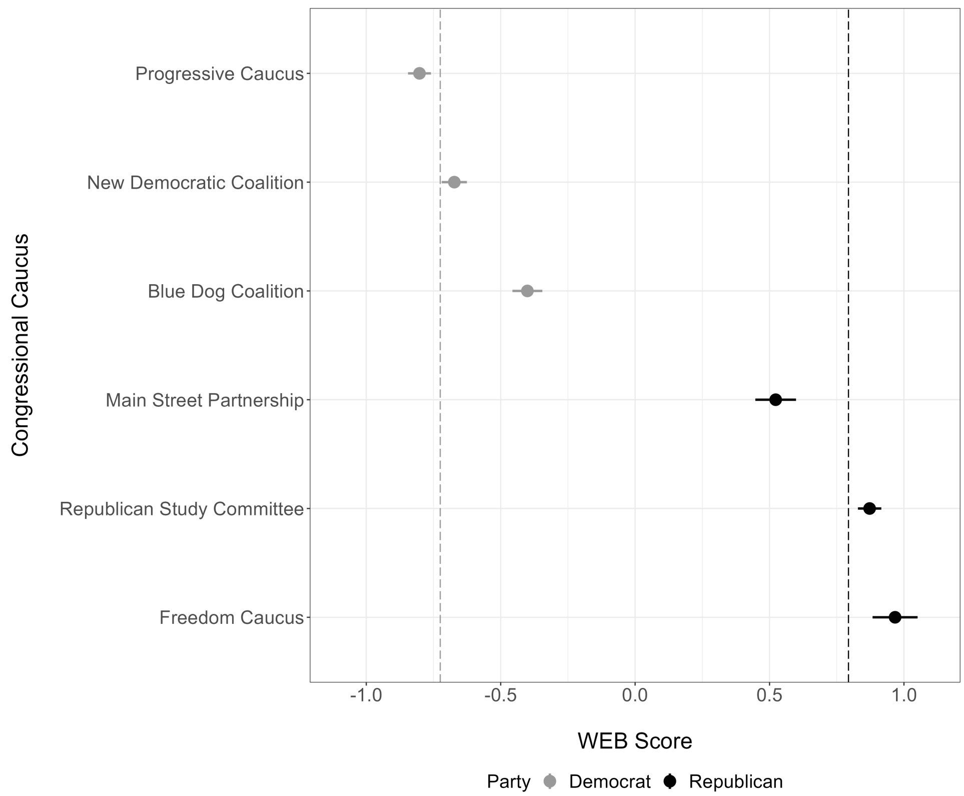
Figure 5 plots the mean WEB Score by caucus for incumbent candidates running in 2018, 2020, and 2022, as well as 95% confidence intervals. In addition, vertical dashed lines depict the mean WEB Scores for incumbent candidates in the Democratic and Republican parties, respectively. Starting with the Democratic Party, WEB Scores pick up on intra-party differences by caucus membership. Incumbent candidates in the Progressive Caucus have the lowest score at an average of -0.80. This is less than the New Democratic Coalition, which has a mean of -0.67 (diff $= -0.13$, p-value ≤ 0.001). The New Democratic Coalition has an average value significantly larger than the Blue Dog Coalition, which has a mean WEB Score of -0.40 (diff $= -0.27$, p-value ≤ 0.001). The differences and the ideological caucus ordering match expectations and provide face validity the measurement is picking up on intra-party differences within the Democratic Party.

Turning to Republican incumbent candidates, WEB Scores also pick up on expected differences by caucus within the party. The more moderate Main Street Partnership has a

[11]Newsweek, "Who Is In House Freedom Caucus? Full List of Members After Midterms Results" November 10, 2022

[12]Caucus membership for Freedom Caucus members who served across multiple Congresses is consistent with Clarke (2020).

Figure 5: Average WEB Score by Caucus Membership



Note: Figure plots the mean WEB Score for incumbent candidates by ideological caucus with 95% confidence intervals. Ideological caucuses are ordered on the y-axis from liberal (top) to conservative (bottom) according to (Clarke 2020). The average position score of all ideological caucuses are statistically different from one another at the $p<0.05$ level.

mean of 0.52. Both the Republican Study Committee, with a mean of 0.87 (diff = 0.25, p-value $\leq$ 0.001), and the Freedom Caucus, with a mean of 0.97 (diff = 0.45, p-value $\leq$ 0.001), have average WEB Scores greater than the Main Street Partnership. WEB Scores also pick up on differences between the Republican Study Committee and the Freedom Caucus, with the Freedom Caucus having a higher average value (diff = 0.10, p-value $\leq$ 0.05). Within both parties, the differences in mean caucus scores provide face validity the measure picks up on differences in candidate positioning within parties.

Internal Validity

One of the advantages of word embedding models is the ability to uncover semantic relationships between words using arithmetic. In the classic example from Mikolov et al. (2013), the authors are able to show:

$$vector[\text{“}king\text{”}] - vector[\text{“}man\text{”}] + vector[\text{“}women\text{”}] = vector[\text{“}queen\text{”}]$$

The ability to uncover these types of semantic relationships between words makes it possible to test a variety of word relationships that should be related to candidate positioning, thus validating the measure against the underlying text. This is possible because word embeddings and candidate embeddings exist in the same dimensional space. As such, if WEB Scores are capturing variation in candidate positioning, they should also be related to certain semantic relationships. Take the following example between the word “universal” and the word “healthcare.” Given that advocating for universal healthcare is more often done by liberal candidates, it should be expected that the relationship between these words is closer for liberal candidates (e.g., Alexandria Ocasio-Cortez (D-NY)) than with conservative candidates (e.g., Chip Roy (R-TX)). This comparison can be done by adding the candidate embedding to the word embedding for healthcare and then assessing the cosine similarity between the new embedding and the word embedding for “universal.”[13] It should be expected this similarity is greater for the more liberal candidate. Specifically:

$$cosine(vector[\text{“}healthcare\text{”}] + vector[\text{“}Ocasio - Cortez2022\text{”}], vector[\text{“}universal\text{”}]) \geq$$
$$cosine(vector[\text{“}healthcare\text{”}] + vector[\text{“}Roy2022\text{”}], vector[\text{“}universal\text{”}])$$

As expected, the cosine similarity for Ocasio-Cortez is 0.31 versus 0.15 for Roy, showing the semantic similarity between “universal” and “healthcare” is closer for Ocasio-Cortez than it is for Roy.

To more formally carry out this test, I rely on the notion that $vector[candidate] + vector[policy]$ should be more similar to a conservative (liberal) policy proposal embedding across candidates as WEB Scores increase (decrease). In developing policy proposal

[13]Cosine similarity assesses the angle between the two vectors. This method is ideal for capturing vector similarities in a high number of dimensions.

Table 4: Policy Word, Policy Stances, and Keywords for Internal Validity Test

Policy Area	**Policy Stance**	**Keywords**
Abortion (Heritage)	Banning abortions after fetal heartbeat	prolife, families, heartbeat, prohibit
Education (Heritage)	Increasing parental involvement in curriculum	parents, choice, homeschooling, transparency
Government (Heritage)	Reducing government spending and regulation	spending, regulations, prudent, fiscal
Immigration (Heritage)	Increasing border security	incursions, enforces, prosecutes, secures
Environment (Justice Democrats)	Increasing renewable energy sources and protecting vulnerable communities	renewable, climate, fossil, color
Guns (Justice Democrats)	Increasing gun control	background, ban, assault, safety
Healthcare (Justice Democrats)	Implementing single-payer health insurance	universal, singlepayer, expand, medicareforall
Wages (Justice Democrats)	Increasing the minimum wage	living, minimum, affordable, cost

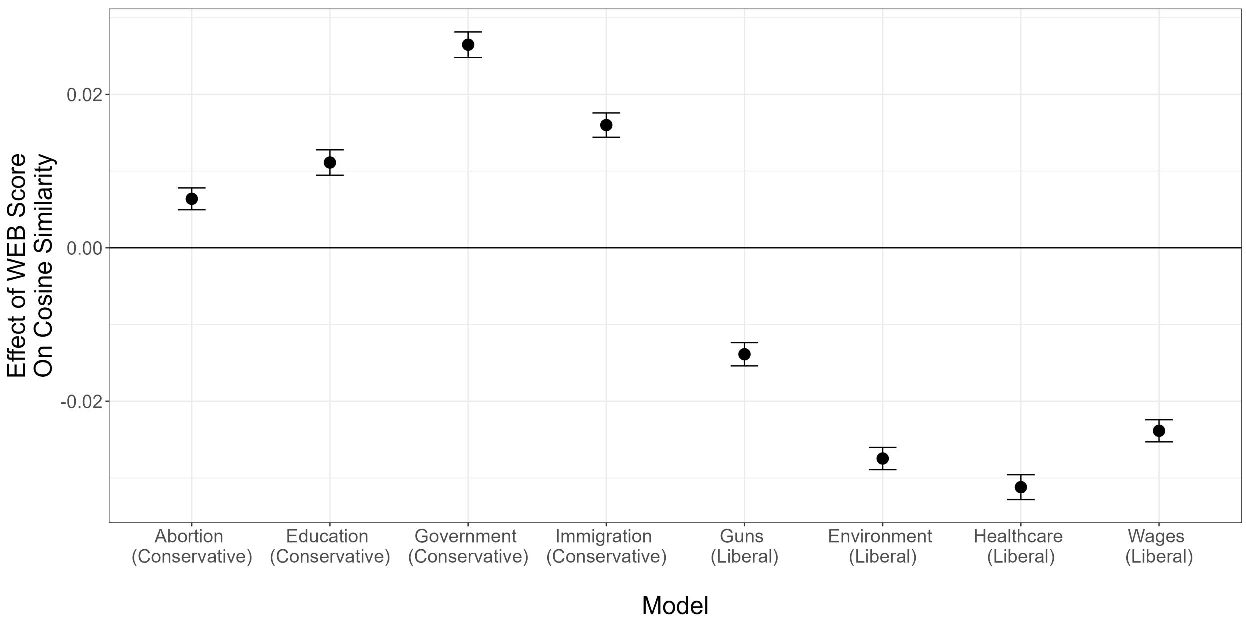
embeddings, I rely on Distributed Dictionary Representations (DDR; Garten et al. 2018). The advantage of this method is that by averaging word embeddings, it is possible to capture a distinct psychological construct. For the purposes of this paper, I use DDR to develop average policy position embeddings that can conceivably be classified as either more liberal or more conservative. To do so, I rely on a set of eight anchoring vignettes that represent the end points of the position scales – four from the Justice Democrats Policy Priorities in 2022, and four from the Heritage Foundations Policy Priorities in 2022.[14]

From each of the policy priorities, I select a set of keywords that are present in the stance each organization is taking. The policy word, policy stance, and policy proposal keywords can be found in Table 4. Full issue statement vignettes can be found in Appendix C.[15] To carry

[14]The Justice Democrats and Heritage Foundation are chosen because they lay out clear, detailed policy positions and self-describe as placing themselves at the extreme of the positioning scale. Justice Democrats outline their mission "is to build a mission-driven caucus in Congress by electing more leaders like Alexandria Ocasio-Cortez and Jamaal Bowman, who will represent our communities in Congress and fight for bold, progressive solutions to our current crises." The Heritage Foundation states their mission is to "formulate and promote public policies based on the principles of free enterprise, limited government, individual freedom, traditional American values, and a strong national defense." This provides face validity to the anchoring vignettes.

[15]One of the advantages of embeddings, and DDR specifically, is that not all words need to be included in the dictionary. For example, because regulation and regulations are syntactic pairs, the inclusion of both

Figure 6: Effect of Candidate WEB Scores on Policy Cosine Similarities



Note: Figure presents coefficient estimates and 95% confidence intervals from the effect of WEB Scores on cosine similarities for each candidate and the relevant policy area. Full regression output can be found in Appendix D.

out the test, I add each candidate embedding to the word embedding for each policy area (e.g., government). I calculate an average embedding of the keywords, and calculate cosine similarities between the resulting candidate-policy embedding and the keyword embedding for each candidate in each policy area.

After calculating the relevant cosine similarities, I conduct eight OLS regressions where the dependent variable is the cosine similarity for each policy area and the independent variable is candidates' WEB Score. If WEB scores are picking up on important semantic relationships related to candidate positioning, it should be expected that the coefficient is positive (negative) for conservative (liberal) policies.

Figure 6 plots the coefficient for WEB Scores from all eight models.[16] Conservative policies are on the left side of the table and liberal policies are on the right side of the table. Across the four conservative policies, the effect of WEB Scores is positive. This means that as WEB Scores increase, the cosine similarity between $vector[candidate] + vector[policy]$ and

adds little to the set of keywords.

[16]Full regression tables can be found in Appendix D.

the average of $vector[keywords]$ for each policy area increase. This can be interpreted as the word embedding for policy words and the average word embedding for the keywords as being more similar for conservative candidates than liberal candidates. For the liberal policies, the effect is negative, as expected. This means as WEB Scores decrease, the cosine similarity between $vector[candidate] + vector[policy]$ and the average of $vector[keywords]$ for each policy area increase. These results provide further validation the measure is picking up on positional differences across candidates.

Application

As discussed previously, one of the main advantages of WEB Scores is the expanded coverage of candidates. This creates new opportunities to test certain research questions that had not been feasible before. One such question relates to the earlier example of Congressman Andy Kim and Reuvan Hendler: how do incumbent candidates position themselves in response to extreme primary challengers? With the increase in extreme challengers in primary elections, research has identified a number of different ways incumbent candidates change their behavior in response. For example, Jewitt and Treul (2019) find that incumbent candidates challenged by an extreme primary challenger vote with their party less often once in Congress. Other work highlights how incumbents change what media sources they share in response to an electoral challenger (Macdonald et al. 2022). Prior work has not tested the relationship between challenger positioning and incumbent positioning due to data limitations. If candidates are trying to maximize their chances in primary elections, it should be expected that they respond to the position of the challenger in the election, and become more extreme (moderate) in their issue positioning if their challenger is extreme (moderate) (Downs 1957).

WEB Scores are well-suited for a research question of this type, given that the most likely candidates to challenge incumbents in primary elections are candidates without prior elected experience (Jacobson 1989; Porter and Treul 2023). To give an idea of the coverage increase, in 2018, 2020, 2022, 462 (48.1%) incumbents faced a challenger in the primary election.[17] Among those primary challengers, 334 (72.3%) have a WEB Score. For comparison, CFscores, which provide the highest coverage of these challengers compared with other measures, only include 121 challengers (41.0%) in 2018 and 2020.[18] This represents a significant drop-off in the number of cases included.

I rely on two measurement approaches to test how incumbent candidates respond to primary election challengers. In both cases, the dependent variable, incumbent position extremity, is measured as an incumbent candidate's WEB Score minus their party's average WEB Score. I then multiply the Democratic candidates' score by -1 to provide a consistent measure across parties. Therefore, positive (negative) values are interpreted as candidates becoming more extreme (moderate).

For the key independent variable of interest, the positioning of a primary challenger, I focus on only the challenger with the highest vote share. I employ both a dichotomous and a continuous approach. In the dichotomous approach, I include a three level factor variable for whether the incumbent was not challenged (omitted category), was challenged by an extreme candidate, or was challenged by a moderate candidate. I classify candidates as extreme if they had a WEB Score greater than their parties' mean WEB Score, and moderate otherwise. For the continuous approach, I measure the challenger positioning in the same way I measure the dependent variable: I subtract the challenger's party mean WEB Score from the challenger's own WEB Score. I then multiply Democratic candidates' measure by -1 for a consistent measure across parties. It should be noted, that in the continuous approach, because challenger positioning is not observed when a challenger does not emerge, this model only includes candidates who are challenged in the primary. The dichotomous approach includes all incumbents.

In addition to the measure for challenger extremity, I also include two models with incumbent-level fixed effects and two models with incumbent-level random effects. Because

[17] This excludes candidates in top-two, top-four, and jungle primary states.

[18] For direct comparison, 202 (68.4%) of candidates who are challenging incumbents in 2018 and 2020 have WEB Scores.

Table 5: Incumbent Extremity and Challenger Positioning

	Dependent variable:			
	Incumbent Extremity			
	(1)	(2)	(3)	(4)
Extreme Challenger	0.057**	0.049		
ref: No Challenger	(0.028)	(0.033)		
Moderate Challenger	−0.033	−0.009		
ref: No Challenger	(0.033)	(0.039)		
Challenger Extremity			0.141***	0.083
			(0.046)	(0.071)
Constant	0.023	0.378	0.048	0.369
	(0.025)	(0.241)	(0.039)	(0.244)
Observations	795	795	315	315
Candidate Random Effects	✓		✓	
Candidate Fixed Effects		✓		✓
Year Fixed Effects	✓	✓	✓	✓

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

most incumbents over the period of study run in multiple elections, the fixed and random effects by incumbent account for unobserved heterogeneity that could affect a candidate's positioning. In essence, both the fixed and random effects by incumbent account for baseline levels of a candidate's positioning that could be affected by district characteristics, primary type, or candidate characteristics such as gender or seniority, among other factors.[19] As a result, the coefficient can be interpreted as capturing within-incumbent variation as a result of changes in the status and positioning of a primary challenger. I also include year-level fixed effects in all models.

Table 5 presents the result of both measurement approaches. The first two columns include all incumbents running in partisan primaries in 2018, 2020, and 2022 using the di-

[19]One limitation of this approach is changes in some candidates' district in 2022 due to redistricting. Because this would not be captured by the fixed and random effects by candidate, I also estimate this same model in Appendix E with only the 2018 and 2020 elections. The substantive results across all models are the same.

chotomous variable approach specified above. The last two columns include only incumbents who were challenged in a primary election and use the continuous measurement approach. Columns 1 and 3 use candidate random effects, and columns 2 and 4 use candidate fixed effects.

Starting with the dichotomous variable approach in columns 1 and 2, the results provide support that facing an extreme challenger (compared with no challenger) is associated with an increase in the extremity of the incumbent candidate. This effect is significant in the random effects model (p-value<0.05) but not in the fixed effects model (p-value=0.139). There is no effect of a moderate challenger on the positioning of the incumbent when compared with no challenger in the race. Turning to the continuous approach in columns 3 and 4, the positioning of the challenger is again associated with a change in incumbent behavior: as the positioning of the challenger becomes more extreme, the positioning of the incumbent does as well. This effect is again significant in the random effects model (p-value<0.01) but not the fixed effects model (p-value=0.248). The results here provide support for a Downsian theory of candidate positioning in primary elections in response to extreme challengers.

Conclusion

This paper introduces Website EmBedding (WEB) Strategic Positioning Scores which improves upon the limitations of prior measurements of strategic candidate positioning. Namely, it increases the number of candidates included and is based on underlying data that actually captures candidate positioning. In addition, the measure possesses high construct validity, both with other aspects of candidate positioning and with the underlying campaign text. The benefits of this new measure, as well as the word and candidate embeddings, expand the number of substanative research questions that can be answered as it relates to candidate positioning. For example, as I show above, WEB Scores are well-suited to assess the extent to which the extremity of primary challengers is associated with an increase in the extremity of incumbent candidates. This result provides further substantive understanding to the role primary elections play to increasing elite polarization in the United States.

As discussed above, scholars have taken a number of different approaches to measure candidate positioning, as well as related concepts. Scholars of congressional elections need to be judicious about their choice of measurement when it comes to candidate positioning, regardless of its use as a focal point or a model control. There are a number of factors researchers should take into account when picking a measure of positioning. As a starting point, researchers need to consider their conceptual quantity of interest. Existing measures cover a variety of different concepts, from voting behavior of candidates both in (e.g., Poole and Rosenthal 1985) and out (e.g., Shor and McCarty 2011) of Congress, to perceptions of candidates' positioning (e.g., Bonica 2014; Christopher et al. 2015), to the actual issue positions of candidates (WEB Scores). All three are substantively important concepts, and researchers must consider what concepts they are primarily interested in capturing. For example, Tausanovitch and Warshaw (2017) note that several proxy measures of candidate positioning provide little value in predicting future legislative behavior in Congress; WEB Scores are not designed to predict this future legislative behavior and should not be used as such. When it does, however, come to actual candidate issue positioning, WEB Scores are uniquely situated to capture this construct.

Second, researchers need to take into consideration the population of candidates they are interested in observing. Compared with previous measures, WEB Scores provide more candidate coverage, especially for inexperienced candidates and candidates in primary elections. One explicit limitation of WEB Scores is the broader scope of the measure. While Porter, Treul and Case (2023) have collected comprehensive data for 2018, 2020, and 2022, the data collection does not exist prior to 2018. Researchers interested in observing longer longitudinal data would be better suited to pick a measure such as CFscores. In addition, WEB Scores cannot be calculated for potential candidates who may run for Congress. If researchers are interested in studying candidate emergence, both recipient CFscores and contributor CFscores provide more leverage in studying those who do not actually end up running for office.

Finally, researchers need to consider their theoretical argument when choosing a measure of candidate positioning. There are a range of research questions where, if the measure approach were based on the perception of candidates' positions, the measure could be endogenous. In these instances, it may be the case that citizens' perceptions are not only based on the candidates' actual positioning, but other related behaviors of interest. When this occurs, it is important to choose a measure that is capturing actual candidate behavior, like WEB Scores, instead of a perception-based measure.

References

Ahler, Douglas J, Jack Citrin and Gabriel S Lenz. 2016. "Do Open Primaries Improve Representation? An Experimental Test of California's 2012 Top-Two Primary." *Legislative Studies Quarterly* 41(2):237–268.

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

Barber, Michael. 2022. "Comparing campaign finance and vote-based measures of ideology." *The Journal of Politics* 84(1):613–619.

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

Black, Duncan. 1948. "On the rationale of group decision-making." *Journal of political economy* 56(1):23–34.

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

Case, Colin R. 2023. "Coloring within the Party Lines: Candidate Branding in Primary Elections.".

Christopher, Hare, Bakker Ryan, Carroll Royce et al. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59(3):759–774.

Clarke, Andrew J. 2020. "Party Sub-Brands and American Party Factions." *American Journal of Political Science* 64(3):452–470.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The statistical analysis of roll call data." *American Political Science Review* 98(2):355–370.

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.

Duck-Mayr, JBrandon and Jacob Montgomery. 2022. “Ends against the middle: Measuring latent traits when opposites respond the same way for antithetical reasons.” *Political Analysis* pp. 1–20.

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.

Gaynor, SoRelle W, Kristina Miler, Pranav Goel, Alexander M Hoyle and Philip Resnik. 2022. “Do You Walk the Walk, Talk the Talk, or Tweet the Tweet? Legislators’ Ideal Points Across Venues.”.

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 behaviour* 6(7):975–987.

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

Hall, Andrew B. and James M. Snyder. 2015. “Candidate Ideology and Electoral Success.”.

Herrnson, Paul S, Costas Panagopoulos and Kendall L Bailey. 2019. *Congressional elections: Campaigning at home and in Washington*. Cq Press.

Hirano, Shigeo, Gabriel S Lenz, Maksim Pinkovskiy and James M Snyder Jr. 2015. “Voter learning in state primary elections.” *American Journal of Political Science* 59(1):91–108.

Jacobson, Gary C. 1989. "Strategic Politicians and the Dynamics of U.S. House Elections, 1946-1986." *American Political Science Review* 83(3):773–93.

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

Jewitt, Caitlin E and Sarah A Treul. 2019. Ideological primary competition and congressional behavior. In *Congress & the Presidency*. Vol. 46 Taylor & Francis pp. 471–494.

Lau, Jey Han and Timothy Baldwin. 2016. "An empirical evaluation of doc2vec with practical insights into document embedding generation." *arXiv preprint arXiv:1607.05368* .

Laver, Michael, Kenneth Benoit and John Garry. 2003. "Extracting policy positions from political texts using words as data." *American political science review* 97(2):311–331.

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

Macdonald, Maggie, Megan A Brown, Joshua A Tucker and Jonathan Nagler. 2022. "To Moderate, Or Not to Moderate: Strategic Domain Sharing by Congressional Campaigns." *Or Not to Moderate: Strategic Domain Sharing by Congressional Campaigns (April 5, 2022)* .

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado and Jeffrey Dean. 2013. "Distributed Representations of Words and Phrases and their Compositionality." *Advances in neural information processing systems* .

Mikolov, Tomáš, Wen-tau Yih and Geoffrey Zweig. 2013. Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies*. pp. 746–751.

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.

Nyhan, Brendan and Jacob M Montgomery. 2015. "Connecting the candidates: Consultant networks and the diffusion of campaign strategy in American congressional elections." *American Journal of Political Science* 59(2):292–308.

Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29(2):357–384.

Porter, Rachel, Maura McDonald and Sarah Treul. 2021. "Changing the Dialogue: Descriptive Candidacies and Position Taking in Campaigns for the US House of Representatives.".

Porter, Rachel and Sarah A. Treul. 2023. "Evaluating (In)Experience in Congressional Elections." `https://rachelporter.org/files/inexperience.pdf`.

Porter, Rachel, Sarah A. Treul and Colin R. Case. 2023. "Database on Primary Election Website Content." *Chapel Hill, NC; University of North Carolina Libraries* .

Ramey, Adam. 2016. "Vox populi, vox dei? Crowdsourced ideal point estimation." *The Journal of Politics* 78(1):281–295.

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

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

Rodriguez, Pedro L, Arthur Spirling and Brandon M Stewart. 2021. "Embedding Regression: Models for Context-Specific Description and Inference." *American Political Science Review* pp. 1–20.

Satopaa, Ville, Jeannie Albrecht, David Irwin and Barath Raghavan. 2011. Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*. IEEE pp. 166–171.

Shor, Boris and Nolan McCarty. 2011. "The ideological mapping of American legislatures." *American Political Science Review* 105(3):530–551.

Slapin, Jonathan B and Sven-Oliver Proksch. 2008. "A scaling model for estimating time-series party positions from texts." *American Journal of Political Science* 52(3):705–722.

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

Tausanovitch, Chris and Christopher Warshaw. 2017. "Estimating candidates' political orientation in a polarized congress." *Political Analysis* 25(2):167–187.

Thomsen, Danielle M. 2014. "Ideological Moderates Won't Run: How Party Fit Matters for Partisan Polarization in Congress." *Journal of Politics* 76(3).

Thomsen, Danielle M and Andrew Hall. 2018. "Who Punishes Extremist Nominees? Candidate Ideology and Turning Out the Base in US Elections." *American Political Science Review* .

Treul, Sarah A and Eric R Hansen. 2023. "Primary Barriers to Working Class Representation." *Political Research Quarterly* p. 10659129231154914.

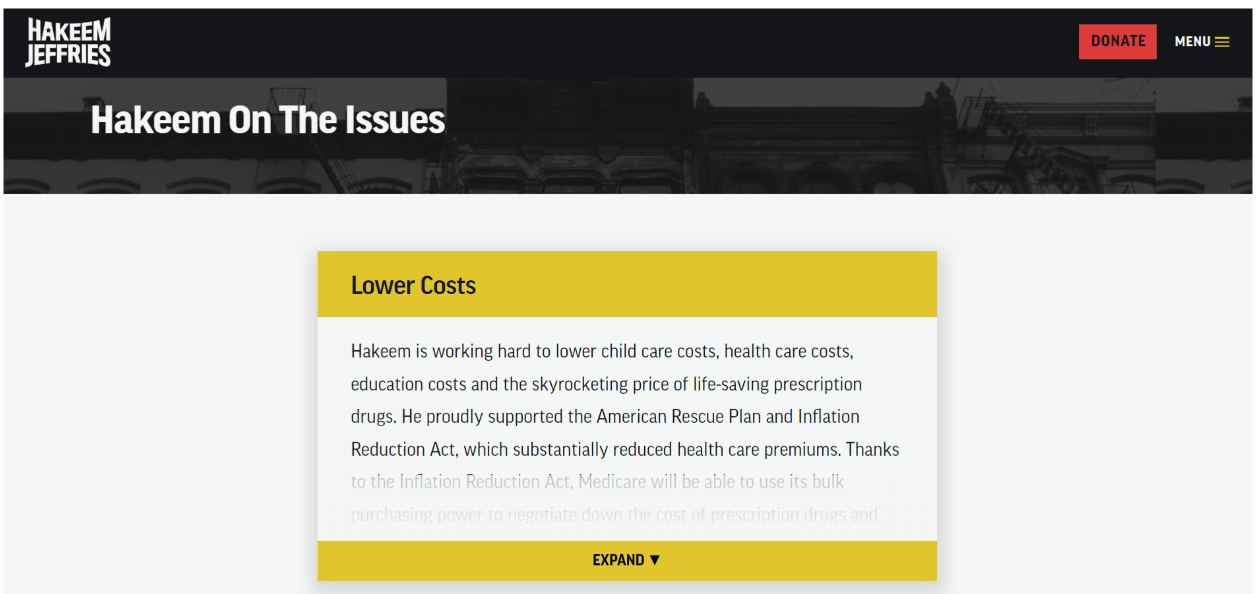
Vafa, Keyon, Suresh Naidu and David M Blei. 2020. "Text-based ideal points." *arXiv preprint arXiv:2005.04232* .

Supporting Information Table of Contents

Appendices

A Example Website Issue Pages

Figure 1: Examples of Campaign Issue Pages from 2022 Congressional Primary Candidates



Note: The top image is from Rep. Dan Crenshaw (R-TX) and the bottom image is from Rep. Hakeem Jeffries (D-NY).

B Correlation Tables for Different Model Parameters

Table 1: Candidate Positioning Correlation Table with Different Model Parameters (window size, embedding dimension)

	5, 100	5, 200	5, 300	6, 100	6, 200	6, 300	7, 100	7, 200	7, 300	8, 100	8, 200
5, 200	0.995										
5, 300	0.994	0.996									
6, 100	0.997	0.995	0.994								
6, 200	0.995	0.997	0.996	0.995							
6, 300	0.993	0.996	0.997	0.993	0.996						
7, 100	0.997	0.994	0.993	0.997	0.994	0.993					
7, 200	0.994	0.997	0.996	0.995	0.997	0.996	0.994				
7, 300	0.993	0.996	0.996	0.993	0.996	0.997	0.993	0.996			
8, 100	0.996	0.994	0.992	0.997	0.994	0.993	0.997	0.994	0.993		
8, 200	0.994	0.996	0.995	0.995	0.997	0.996	0.995	0.997	0.996	0.994	
8, 300	0.992	0.995	0.996	0.993	0.996	0.997	0.993	0.996	0.997	0.993	0.996

C Full Issue Statements for Internal Validity Test

Table 2: Full Issue Statements from Heritage Foundations

Policy Area	Issue Statement
Abortion	All children conceived deserve to be born to married mothers and fathers who will love, guide, and protect them throughout their lives, but **family** breakdown and rampant abortion have torn apart the soul of our country and sapped it of its strength and moral authority. We will advance the **Heartbeat** Protection Act to **prohibit** abortion nationwide after the moment a heartbeat can be detected. At the state level, we will work with governors, legislators, and other state-based allies to pass heartbeat laws (or better) on abortion. We will work to **prohibit** the interstate commerce of abortion pills in **pro-life** states by advancing legislation in both the House and Senate.
Education	The Heritage enterprise will work to minimize the federal government's intervention in education. The education system is failing our children—from the scourge of woke ideas like critical race theory and radical gender ideology to the lack of accountability to **parents** and an absence of academic **transparency**. **Parents**, not bureaucrats, should be making teaching and learning decisions that align with their values. Taxpayer dollars should help students to succeed with a great education, not prop up failing school systems. The entire Heritage enterprise will spearhead reforms at the state level to protect parental rights and expand education **choice** and will work at the federal level to limit Washington's intervention, ultimately driving a clarion call to eliminate the U.S. Department of Education. Minimizing federal intervention in education includes supporting the introduction of federal legislation to (1) give states more budget authority over federal education funding with fewer strings, (2) reduce federal intervention in early childhood education by reforming programs such as Head Start, and (3) expand families' access to homeschooling by reforming 529 savings accounts to include **homeschooling** expenses and by expanding and making permanent the D.C. Opportunity Scholarship Program.
Government	Government **spending**, **regulations**, and inflation are a tax on all Americans, especially working families who struggle to make ends meet. **Prudent fiscal** decisions by government can enable American families to flourish without politicians and bureaucrats controlling their lives. The Heritage enterprise will advance a blueprint to reduce the size and scope of the federal government, ensure that government spends less of our money to save us from falling off the **fiscal** cliff, and stop the growth of federal **regulations**.
Immigration	Americans should be able to live peacefully without constant fear of crime or **incursions** across our borders. A strong justice system **enforces** existing U.S. law, **prosecutes** criminals, **secures** our borders, and preserves our national identity. America's current border crisis and the level of crime in many cities are out of control, and the human costs are staggering.

Note: Keywords from Table 4 are bolded in each issue statement.

Table 3: Full Issue Statements from Justice Democrats

Policy Area	**Issue Statement**
Environment	Now is the time for a comprehensive, once-in-a-generation mobilization that prioritizes front-line communities, combats the **climate** crisis, and creates millions of good-paying union jobs. A Green New Deal will transition away from **fossil** fuels and dramatically expand existing **renewable** power sources with the goal of meeting 100% of national power demand through **renewable** sources. A Green New Deal also provides people across the country with the opportunity, training and education needed to participate fully and equally in a green economy, offering jobs to help rebuild our crumbling infrastructure. A Green New Deal ensures a just transition for all workers, low-income communities, communities of **color**, and indigenous communities.
Guns	Gun violence is a public health crisis in the United States that disproportionately impacts communities of color. More than 90 percent of Americans support expanded **background** checks, 54 percent want a **ban** on **assault** weapons and 54 percent want a ban on high capacity magazines. We agree with the majority of the American people and support these measures. To enact common sense gun **safety** measures, we must break the NRA's hold on our corrupt government and prioritize the mental and physical health of the people over the billion-dollar gun manufacturing industry's bottom line.
Healthcare	The United States has the most expensive and least effective healthcare system compared with other industrialized nations. It's time to end the destruction of healthcare in America by price gouging, for-profit private health insurers and catch up to every other modern nation that's implemented a **single-payer universal** healthcare system – no networks, no premiums, no co-pays, no deductibles and no surprise bills. **Medicare-For-All** will **expand** Medicare coverage to include dental, hearing, mental health and substance abuse treatment, prescription drugs, long-term and disability care, and reproductive and maternity care. We must also invest in frontline care workers who are the backbone of our economy.
Wages	Over the past several decades, the **cost** of **living** has increased significantly while workers' wages have remained relatively stagnant. While CEO's compensation soars, most workers' wages aren't even keeping up with inflation and **affordable** housing remains out of reach. We must secure a **minimum** wage of at least $15 that's tied to inflation.

Note: Keywords from Table 4 are bolded in each issue statement.

D Internal Validity Test Regression Tables

Table 4: Full Regression Tables from Figure 6 for Conservative Policies

	Dependent variable:			
	Government	Immigration	Abortion	Education
	(1)	(2)	(3)	(4)
WEB Scores	0.026***	0.016***	0.006***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.261***	0.326***	0.366***	0.310***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	4,554	4,554	4,554	4,554
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5: Full Regression Tables from Figure 6 for Liberal Policies

	Dependent variable:			
	Environment	Healthcare	Guns	Wages
	(1)	(2)	(3)	(4)
WEB Scores	−0.027***	−0.031***	−0.014***	−0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.325***	0.262***	0.399***	0.512***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	4,554	4,554	4,554	4,554
Note:				*p<0.1; **p<0.05; ***p<0.01

E Table 5 Replication with 2018 and 2020

Table 6: Incumbent Extremity and Challenger Positioning: 2018 and 2020 Elections

	Dependent variable:			
	Incumbent Extremity			
	(1)	(2)	(3)	(4)
Extreme Challenger	0.078**	0.074		
Ref: No Challenger	(0.037)	(0.050)		
Moderate Challenger	−0.098**	−0.117**		
Ref: No Challenger	(0.044)	(0.055)		
Challenger Extremity			0.177***	0.142
			(0.057)	(0.110)
Constant	0.019	−0.060	0.052	0.250
	(0.027)	(0.162)	(0.040)	(0.235)
Observations	518	518	189	189
Candidate Random Effects	✓		✓	
Candidate Fixed Effects		✓		✓
Year Fixed Effects	✓	✓	✓	✓

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