

Running as a Woman?: Candidate Presentation in the 2018 Midterms

Maura McDonald
mgm92@live.unc.edu

Rachel Porter
rachsur@live.unc.edu

Sarah Treul
streul@unc.edu

Department of Political Science
University of North Carolina at Chapel Hill

November 8, 2020

Abstract

The record high number of women who ran for the U.S. Congress during the midterm elections led many journalists proclaim 2018 as another “Year of the Woman.” Although not every female candidate was successful, this large number of women running for office provides the opportunity to advance our understanding of the ways in which women present themselves to their voters. Using the “Biography” pages of over 1,500 2018 congressional campaign websites, we utilize a structural topic model to examine how these candidates present themselves to their constituencies. In doing this, we find great variance in the presentation styles of women running for Congress in 2018. We also find that prior political experience, more so than gender, is the primary driver in influencing how candidates (both men and women) present themselves. Experienced candidates utilize similar styles that highlight their past political work while amateur candidates are more likely to use “values-driven” language.

Introduction

The 2018 midterm elections marked a historic occasion for female representation in the United States. Since 1992, a mere 160 female candidates, on average, ran for the House in any given election cycle. However, in 2018, a record-breaking 476 women ran for the House of Representatives, and 36 women were newly elected to office. The year was also unique as women's issues, such as gender equality and workplace harassment, were frequently at the forefront spurred by the Women's March and the #MeToo movement. Taken together, this increase in female candidate emergence and the centrality of gender issues on the national stage provide us with a unique opportunity to study the role of gender in how a candidate presents herself to voters (i.e. the candidate's "presentation of self" (?)). In particular, the presence of a large number of female candidates allows us to examine how women may differ, not only from men, but also from other women as they develop a style that determines how they tailor communication towards their voters and potential constituents. Emphasizing one's gender by running "as a woman" seems like a highly advantaged strategy in a year that many called another "year of the woman." Therefore, if there ever was an election to expect pronounced differences in gendered self-presentation, it would be in this election.

An increasingly crowded field of women illustrates the necessity for women to develop distinctive self-presentation strategies as a way to distinguish themselves from one another and to attract support from voters, donors, and party elites. This is particularly true in primary elections, where voters are not able to use party as a cue. This need for distinctiveness raises important questions about what candidates choose to emphasize and deemphasize over the course of a campaign. Do women choose to highlight their gender in similar ways, creating a homogenous presentation style that is unique to female candidates? Or, given that many of the factors that traditionally influence elections do so "regardless of the sex of the candidates" (?), is it more likely that women, like men, are heterogeneous in their presentation styles? Should we then expect to see more

homogeneity between men and women who share other characteristics (i.e. Democratic women are more similar to Democratic men than they are to Republican women). If so, how do differences in party and previous experience in elected office influence what candidates choose to emphasize?

To answer these questions, we turn to campaign websites. Throughout the course of 2018, we collected the text from the “About” pages of all candidates who ran in the 2018 primary elections—totaling 1,510 candidate biographies of which 24% were women.¹ The “About” pages of candidate websites are the ideal place to examine self-presentation. According to ?, p.345, candidate campaign websites are a uniquely ideal form of data for studying campaign communication because they are “unmediated (i.e., directly from the campaign), complete (i.e. covering a full range of rhetorical strategies), and representative of the population of campaigns.” Additionally, ? find that the typical candidate presents a campaign agenda on her website that covers nearly twice as many issues when compared to her advertisements, indicating that these websites contain a more comprehensive view of a candidate’s positions, containing information that likely shows up throughout the campaign. Since candidates are not limited to the same time and space restrictions they might be in campaign advertisements, press releases, or social media posts, they can emphasize everything that they think might be important to potential supporters, including copies of advertisements or links to speeches or articles (??). Though the audiences for these websites are more homogeneous than those for campaign ads (?), we expect that candidates use these websites to present a strategic picture of themselves. The “About” page of a candidate’s website is a picture of the candidate as she wishes to be seen by voters and how she hopes the media will portray her.

Thus, the campaign websites provide us with an unmediated snapshot of the candidate, or “the aggregation of campaign communications that reflect a campaign’s overall rhetorical strategy” (?, p.345). In this paper, we specifically focus on the candidate’s

¹This number includes only candidates who had a campaign website, as we could not use an automated approach to collect off of Facebook or LinkedIn.

biographies or “About” pages.² These pages are more than just a place for candidates to paste their resumes. Candidates use these pages to make a case for why they are the best for the job—we think of them as “Why Me” pages. These “Why Me” pages are a place where “different aspects of personal history are adjusted in accordance with a strategic plan” making them an ideal place to investigate how candidates piece together different elements of their campaign to make the most compelling case to voters for their election. (?, p.281).

Gender, Experience, and Party in Political Campaigns

In order to understand the decisions women make regarding the presentation of their gender, it is first necessary to understand the role that gender plays in modern day elections. Though women make up 51% of the national population, they still are significantly underrepresented in Congress. Much of the past literature surrounding female candidate emergence and success suggests that this is not because women cannot be successful at the ballot box. On the contrary, men and women are elected at comparable rates.³ What is likely more responsible for the dearth of women running is that women are more likely to wait to run until they have accrued a substantial amount of experience in elected office, passing some “quality threshold,” often making them more qualified than their male counterparts once they decide to run (?). Reinforcing this theory, ? demonstrate that women are even more likely to hold elected office before seeking a House seat, showcasing their perceived need for previous experience before running for Congress. Relatedly, ? argue that women are underrepresented because they do not run as often as men do, in part because they are less likely to receive early encouragement to run for office and are more likely to put off running due to family considerations. In summary, there are fewer women running because they, as a group, are unconvinced that they are ready to run—let

²A sample of these pages can be found later in the paper.

³For a review of the literature about the history of women’s electoral fortunes see ? and ?.

alone ready to win.

This personal decision making, combined with an uneven electoral playing field, lead women to believe that they must be extraordinarily qualified candidates in order to run for Congress. As a result, women, regardless of previous elected office experience, party, or race type, are likely to focus on and play up their experience, in whatever form they have, as they tend to have a lot of it and it provides the foundation for their sense of qualification. It is here that we expect to see a gendered difference in campaign presentation. Women will be more likely than men to emphasize their political and other types of experience as a function of proving their qualifications to their potential voters and donors.

Turning to the campaign environment, many studies find that women receive unequal media coverage, with more attention paid to their appearance and personality traits than their qualifications (??). This likely encourages women to play up their experience even more in order to counterbalance this uncontrollable media coverage. Additionally, women must be more strategic about their tone, receiving advice to avoid emotion, negative ads, and other strategies that will help them minimize the potentially negative effects of gender stereotypes (??). On the other hand, several studies have also found that women have a distinct advantage when they choose to run “as women” by focusing on issues and traits that voters associate with female candidates (????).

Increasingly, it is also possible for women to incorporate both masculine and feminine stereotypes into their self-presentation (?). However, female candidates still face certain disadvantages when deciding which traits to focus on. If a candidate is seen as too feminine, she risks seeming weak or unqualified, but coming off as too masculine can make her seem calculating or over ambitious (?). Facing this trade-off, past literature suggests that female candidates will likely choose the strategy that best fits their electoral context. That is, women will emphasize their gender if they feel that doing so will appeal to voters or enhance their self-presentation, but if they feel that their gender will detract from their qualifications they are likely to avoid drawing attention to it.

One trait most candidates are willing to emphasize is prior experience in office, as such experience has long been considered a critical part of winning elections (?). More recent work shows that while experience is still a coveted trait, candidates lacking prior experience can still be successful in today's political environment (?). By emphasizing other areas that make them unique, candidates can demonstrate their strengths to voters while drawing attention away from their lack of experience. Therefore, women without previous experience may be particularly likely to play up their gender if they think that it will help build a support base around a more compelling element of their campaign. Though emphasizing electoral qualifications might be a preferred tactic, it is also important for female candidates to play up what makes them distinctive in their race. By positioning themselves as uniquely qualified to speak about and bring attention to the challenges women face (?), candidates will be able to make the case for why they should be the party's candidate, even if they lack legislative experience. Previous research indicates that gender, as well as racial and ethnic identities, play an important role in influencing what elements are emphasized in constituent appeals (????) and, absent experience, we think emphasizing gender as a point of distinction is a smart tactic.

These strategic decisions also likely differ by party. Previous literature shows that gender stereotypes tend to advantage Democratic women more than Republican women (?). This is particularly true in primary elections, where ? find that Republican female candidates have a more difficult time finding success than Democratic women. ? finds that voters who prefer female candidates are more likely to vote for Democratic candidates overall, making them less likely to vote for Republican women. She argues that "voters do not simply want more women in office; they want a particular kind of woman in office" (?, p. 193). Additionally, there is evidence that members of the Democratic Party are more likely to consider diversity as at the core of their party, while Republican coalition members are less likely to embrace "identity-based appeals for the expansion of women in office" (?). Furthermore, ? find that partisan donor pools are more likely to support

Democratic women over Republican women. This indicates that Democratic women will be more likely to play up their gender than Republican women, as it is likely to be more of an asset as it does not detract from their ability to fund their campaign, and meets the expectations of their partisan voters.

Finally, there is increasing evidence that these gender disparities on the campaign trail are diminishing, or may even no longer exist. Several recent studies have found very few gender differences in candidate communication content (???). In today's political environment, party and political experience are more influential in forming voter attitudes and vote choice than gender cues (???). In fact, ? finds that voters are likely to vote for a female candidate, as long as she shares their party. Additionally, while gender disparities certainly have occurred in the past, ?, p.134 find that the "congressional campaign environment is very similar for male and female candidates, and that [what is] holding women back is not discriminatory or more challenging electoral language." As a result, though the literature suggests that there may be some advantages to running "as a woman" and focusing on particular issues or personality traits, it seems likely that playing up these factors is less advantageous if it detracts from key elements of party membership or electoral experience.

Understanding what candidates choose to emphasize and deemphasize depending on these differences is a valuable contribution to determining how candidate communication can influence electoral outcomes. Furthermore, while ?, p.23 find that gender disparities no longer exist in general elections, they suggest that "it is possible that gender could play a stronger role in shaping the dynamics of primaries or non-partisan elections, where voters are unable to rely on partisanship as a decision making shortcut." Here our extensive collection of primary election candidate websites will allow us to examine what candidates of different genders choose to emphasize and deemphasize in primary elections and further extrapolate the ways in which candidates may perceive gender disparities during primary campaigns where they cannot appeal to party alone.

Examining Female Candidate Self-Presentation

Much of the literature has examined the campaign strategies and presentation styles of women compared to men, and recent studies have found few differences between genders. However, very few studies have examined potential heterogeneity between female candidates. This is likely because until 2018 there have been a relatively small number of female candidates in a given election, making it difficult to study. In this way, 2018 is a unique case as the increase in female candidates leads to important variation in experience, and—to a lesser extent—party, that allows us to examine the ways in which women choose different campaign messages across experience, party, and, of course, gender.

Overall, we expect to find that women tailor their biographies based on factors that are important to a candidate’s strategic calculus. In particular, we expect that women may build their “Why Me” pages based upon party and their political experience—not just their gender. Focusing first on experience, as discussed above, the candidate emergence literature finds that women only emerge as candidates when they feel as though they have checked all the boxes and are qualified to hold office. As a result, women are likely to emphasize their experience more so than men in order to prove that they are qualified to hold public office. Because most women tend to wait to emerge as candidates until they feel that they are qualified to hold public office, female candidates often have a lot of experience to advertise. However, not all women who emerge as candidates have political experience or have held elective office. For these inexperienced⁴ women, focusing their attention on other aspects of their background is likely a more successful strategy than focusing on their lack of experience.

Hypothesis 1: Experienced female candidates will emphasize their political experience and qualifications more than their male counterparts.

⁴Here the term “inexperienced” is being used for someone who previously never held elected political office (i.e. a “non-quality candidate” (?)). We also use the word “amateur” to describe someone without prior experience in elected office.

Hypothesis 2: Experienced female candidates will emphasize their political experience and qualifications more than amateur (inexperienced) female candidates.

Beyond differences in experience, we also expect there to be systematic differences in the ways in which female candidates choose to emphasize their gender. In particular, much of the previous literature indicates that being a female candidate is far less of an asset for Republican women than it is for Democratic women (?). As a result, we expect that Democratic women candidates, regardless of their past experience, will be more likely to emphasize their gender than their Republican counterparts because such identity-based appeals are more likely to be well received by Democratic donors and voters. Within the Democratic party, however, we expect to see differences by experience level. We anticipate that amateur—or inexperienced—Democratic women are far more likely to emphasize their gender in their self-presentation strategies than experienced Democratic women. We expect this to be the case because, as noted above, experienced Democratic women are far more likely to emphasize their political experience and will prioritize this trait over emphasizing their gender. Amateur Democratic women, however, will be looking for ways to appeal to voters and may prioritize highlighting their gender as a way to appeal to voters and donors that draws attention away from or provides justification for their lack of experience. This is not to say that we do not expect experienced Democratic women to play up their gender at all. However, we expect that amateur Democratic female candidates will be more likely to emphasize their gender than their experienced counterparts.

Hypothesis 3: Democratic women candidates will emphasize their gender more than Republican women candidates.

Hypothesis 4: Amateur Democratic women candidates will emphasize their gender more than experienced Democratic women candidates.

Data and Methods

To test these hypotheses, we embarked on a substantial data collection effort. We collected the biographical text for all candidates who had a campaign website and ran in a 2018 congressional primary election—amounting to a total of 1,510 candidates. To collect these data, we first identified the names of all major party candidates running in 2018. We then searched for these candidates’ campaign web-pages by following links from Politics1.com, which provided a near complete inventory of sites. We also visited candidate social media pages and conducted a simple Google search in our attempt to identify the complete population of congressional campaign websites. We successfully found campaign websites for 83% of candidates who appeared on a primary ballot in 2018.⁵ Next, we determined whether a given candidate had a biography on her website. For many candidates this was a simple process. Most websites had a dedicated biographical page titled “Meet the Candidate” or “About Me.” For those websites that did not have a dedicated biography page, we defaulted to any descriptive language provided about the candidate on her website’s homepage. Using a mixture of web scraping and hand coding, we pulled the text from these campaign websites for our analysis.

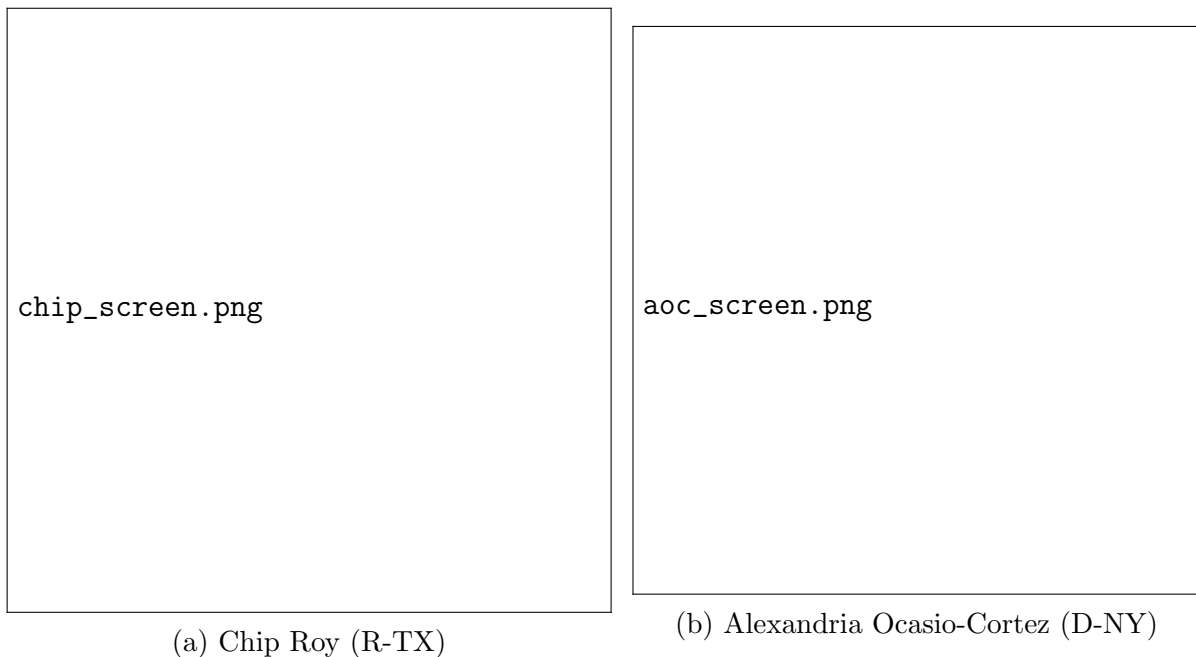
To our knowledge, this is the first data collection effort to harvest text from congressional campaign websites for *all* available candidates running in a given election year. Because of the sheer amount of work involved in compiling and cataloging campaign websites, previous analyses examined only a sample of campaign websites or restricted their scope to the general election (e.g. ??). This makes our data set unique as it provides a near complete picture of the types of candidates who ran for Congress in 2018. We chose to characterize candidate self-presentation using website biographies because, un-

⁵A small number of candidates had no official campaign website. An analysis of the missing websites suggests there are four distinct reasons candidates lack websites: 1) the candidate was an incumbent without any electoral competition (this accounts for 9% of the missing websites); 2) the candidate used Facebook or Twitter as her primary site; 3) the candidate was unsophisticated with little to no campaign presence; and 4) our automated scraper somehow missed the website or there was no biography page as a part of the site (this accounts for 4.5% of the data).

like newspaper profiles or television advertisements, it provides a complete, unmediated narrative of the candidate. Additionally, these web-pages are where the media, people introducing candidates at events, or interviewers are likely to turn for information about the candidates. For these reasons, candidates and their campaign staff spend a substantial amount of time carefully crafting their campaign websites.

Candidates are conscious about their self-presentation; they may play up or play down certain characteristics about themselves based on the district. As previously mentioned, biography pages are a particularly good place to examine candidate self-presentation because these pages go beyond a mere resume. They compile the elements of a candidate’s personal history, experience, and issue positions that are the most salient to the electorate. As a result, these “Why Me” pages are a strategic compilation of information, where a candidate makes the best possible case for herself without the space restraints inherent in other media sources.

Figure 1: Examples of Campaign Biography Pages from the 2018 Midterm Election



Campaign biographies vary from website to website because their content is tailored to win over voters unique to a candidate’s district. However there are also similarities, with most bios featuring information about the candidate’s upbringing, education, notable accomplishments, and connections to the district. Figure 1 depicts two examples of typical campaign biography pages on candidate campaign websites. In Figure 1 (a), Chip Roy (R-TX) emphasizes that he is a “genuine conservative” and “true patriot,” following these characterizations of himself with a list of notable positions in public service, such as Chief of Staff to Ted Cruz (R-TX) and senior advisor to Texas Governor Rick Perry. Alexandria Ocasio-Cortez (D-NY) in Figure 1 (b), on the other hand, begins her biography by establishing her deep roots in the Bronx and emphasizing her Puerto Rican heritage. These contrasting examples demonstrate how candidates tailor their narratives to elevate themselves and their campaigns.

In order to better assess how candidate self-presentation in the 2018 congressional election varied across individual candidate characteristics and race conditions, we linked our text data to meta data on each candidate. Specifically, we employed a primary elections data set by ?. This data set includes standard information on a candidate’s campaign along with data on gender and previous political background or electoral experience. As previously discussed, we posit that in an election year with a record number of female candidates, a candidate’s self-presentation may be defined by more than just her gender. Using these data on candidate characteristics will allow us to test this hypothesis. In order to combine these two data sets into one, we use a probabilistic record linkage model by ? in lieu of unique identifiers for each candidate.

To prepare the text for modeling, we took several pre-processing steps standard in text analysis (?). First, we cleaned the text of any HTML tags and extraneous source code. Second, we removed any stop words — commonly used words such as “the,” “a,” or “in” that have no substantive meaning but rather serve a purely grammatical function. Further, we discarded punctuation, numbers, and removed capitalization. To ensure

we did not discard important information in the form of acronyms or abbreviations we converted these occurrences into full words.⁶ We also chose to remove city and state names.⁷ Third, we simplified biography vocabulary by stemming words, which removes word endings to reduce the dimensionality of text. For instance, using stemming, words like *legislative*, *legislator*, and *legislation* would simplify to *legislat-*. Finally, we removed infrequent words, dropping any terms that did not appear in at least 30 candidate biographies.⁸ After pre-processing, the corpus for our analysis includes 1,510 documents (i.e. the “About” pages we analyzed), 1,584 terms (i.e. the number of *distinct* words across all documents) and 217,899 tokens (i.e. the *total* number of words across all documents).

We use a structural topic model (STM) to gain a better understanding of the content that candidates include in their campaign biographies. Additionally, we use this model to investigate whether the substance of biographies varies based on individual candidate characteristics. At its core, an STM defines topics in a text using an unsupervised machine learning approach. Topics are distributions of semantically cohesive words determined by the topic model and are based on word co-occurrences. Put differently, an STM is able to determine the types of topics or “themes” talked about within a text, grouping words into topics based on how often they are used together. Structural topic modeling is ideal for our purposes because it allows for documents to be expressed as a mixture of topics, rather than just a single topic (?). After reading a sample of campaign biographies, it became clear to us that candidates describe themselves using a variety of personal characteristics and campaign themes. A mixed-membership statistical text model, like an STM, captures this important texture in candidate self-presentation. Furthermore, one of the principle innovations of the STM is its use of covariates in the definition of prior distributions for

⁶For instance, abbreviations like PhD, JD, and MD convey important information about a candidate’s educational background and attainment.

⁷These words do not lower the semantic coherence of our topics in modeling because they do not directly relate to a candidate’s self presentation.

⁸In other words, a term had to appear in at least 2% of our sample of campaign biographies to be included in our analysis. For a full review of documents, words, and tokens removed by threshold see Figure A.1 in Section A.1 of the Appendix.

document-topic proportions and topic-word distributions (?). For our application, this means that attributes like a candidate’s gender or past political experience can be used to help inform where we might expect to see variance in topical content across campaign biographies. The inclusion of covariates in topic modeling is integral to our analysis, allowing us to determine similarities and differences in self-presentation across a variety of cleavages.

Using an STM, we examine the topic prevalence—or how much each topic is discussed—in candidate campaign biographies as a function of gender, past political experience, and partisanship. These three characteristics are included in the structural topic model as an eight-level factor covariate (i.e. experienced Democratic female, amateur Republican male, etc.), which is outlined in Table A.1 of the Appendix. To maximize the number of cases included in each factor level, we grouped incumbents running for reelection with other experienced candidates.⁹ As previously discussed, including this covariate in the STM allows us to compare variability in the prevalence of topics across factor levels. For example, if gender is a central determinant of a candidate’s self-presentation, we would expect to see few differences in topic prevalence across campaign biographies for women Democrats despite differing levels of political experience. We specified twenty-seven topics in this STM, a number chosen using the best practices outlined by ?. A full account of our topic selection strategy can be found in Section A.1 of the Appendix.

The STM topic labels for this model are presented in the first column of Table 1. In order to label the twenty-seven topics, we reviewed the model summary of stems associated with each topic. In determining a topic label, we evaluated in what ways these stems may be similar or how they built toward a common theme. Topic labels were

⁹Because politically experienced candidates and incumbent members of Congress both have elected experience, we expect that these candidates will present themselves similarly. Disaggregating these groups so that incumbents, experienced, and inexperienced are evaluated separately produces weaker but substantively similar results to those in the following section. For the purposes of this analysis, all references to “experienced” candidates also include incumbents running for reelection.

Table 1: Model Topics as a Function of Gender, Party, and Political Experience

Description	Stems	%
Family / Education	school, went, colleg, teach, mother, teacher, father, attend, taught, dad, rais	7.8
Awards / Prestige	award, associ, board, univers, master, bachelor, degre, member, advisori, club	6.7
Representation	peopl, want, believ, need, listen, deserv, better, solut, voic, struggl, understand	6.6
American Dream	hard, dream, small, grow, busi, valu, start, opportun, american, job, owner	6.1
Military	armi, marin, corp, medal, command, air, navi, militari, duti, deploy	5.4
Think-Say-Do	thing, just, go, say, think, even, like, enough, never, tri, someth, lot, get, politician	5.2
Business / Industry	develop, consult, manag, execut, technolog, industri, sector, busi, region, project	4.5
Legislative	committe, caucus, hous, subcommitte, legisl, chair, appropri, congressman	4.5
Tax / Spending	conserv, tax, budget, taxpay, amend, cut, fiscal, spend, govern, regul, limit	4.5
Bipartisanship	improv, infrastructur, strengthen, economi, creat, prioriti, help, across, bipartisan	4.2
Elect. Experience	elect, citi, mayor, district, repres, council, resid, counti, congression, senat	4.1
Healthcare	afford, access, healthcar, qualiti, wage, care, health, equal, medicar, ensur, protect	3.8
Religion	faith, god, church, bless, radio, ministri, book, host, pastor, liberti, man, cultur	3.8
Diversity	women, advoc, neighborhood, organ, student, educ, program, advocaci, commu-niti, divers, congresswoman, obama, nonprofit, latino, champion	3.4
Blue Collar	engin, compani, manag, industri, electr, construct, commerci, truck, mechan	3.3
Activism	progress, corpor, money, climat, worker, power, union, chang, clean, democraci	2.8
President / Party	republican, trump, democrat, parti, agenda, primari, presidenti, candid, navig	2.8
Legal / Lawyer	crimin, attorney, polic, prosecutor, crime, victim, lawyer, judg, legal, investig	2.7
Medicine / Doctor	medic, medicin, doctor, nurs, cancer, patient, health, research, hospit, physician	2.7
Campaign Phrases	run, congress, now, privileg, thank, home, futur, america, time, concern, meet	2.5
Support / Donate	read, donat, pleas, vote, endors, contribut, news, share, email, voter, pledg, event	2.8
Homeland Threat	secur, foreign, terror, discuss, nuclear, defens, obama, threat, terrorist, israel, war	2.2
Agriculture	farm, agricultur, farmer, cross, western, river, rural, fish, valley, blue, lake, bureau	1.8
Coaching	coach, tough, play, scout, sport, team, basebal, boy, footbal, basketbal, littl	2.3
Nature / Land	resourc, face, open, third, sourc, water, servic, famili, land, natur, fourth, light	1.1
Immigration	issu, paid, post, border, divid, status, point, page, color, import, break, name	1.0
Partisan Issues	gun, cost, illeg, requir, drug, million, can, must, control, birth, percent, type	0.9

determined using the top FREX words associated with each topic, which are presented in the second column of Table 1. These FREX words are determined by an equation that weights words by their overall frequency in the text and how exclusive they are to each topic (?). Topic summaries using highest frequency words instead of FREX words can be found in Table A.3 of the Appendix. Using these high frequency stems yields nearly identical topic labels. To validate our topic labeling, we randomly selected and read five campaign biographies that included a relatively high proportion of words associated with each of the twenty-seven topics. The overall prevalence of each topic in campaign website biographies is presented in the third column of Table 1.

Results

We analyze the campaign website biographies for 578 Democratic men, 277 Democratic women, and 90 Republican women.¹⁰ Our sample of candidates is fairly balanced according to gender, with a similar proportion of men and women candidates running in incumbent-held districts and open seats. A detailed breakdown of these descriptive statistics can be found in Table A.2 of the Appendix. A slightly larger proportion of Democratic women have electoral experience compared to Democratic men—31% of Democratic women versus 28% of Democratic men. This is unsurprising given what we know about women preferring to run when they have experience (?).

Using our STM model, we estimate a series of OLS regressions to assess variation in topic prevalence across partisanship, candidate gender, and past political experience. With this approach, we can determine whether there are statistically significant differences in topic prevalence across groups (i.e. experienced Democratic women vs. amateur Democratic women). In these models, the unit of analysis is a candidate biography, the covariate of interest is the partisan-gender-experience factor variable, and the outcome is the proportion of a candidate biography that is about each of the twenty-seven topics

¹⁰This excludes an additional 567 websites of Republican men that we did not analyze here.

described in Table 1.

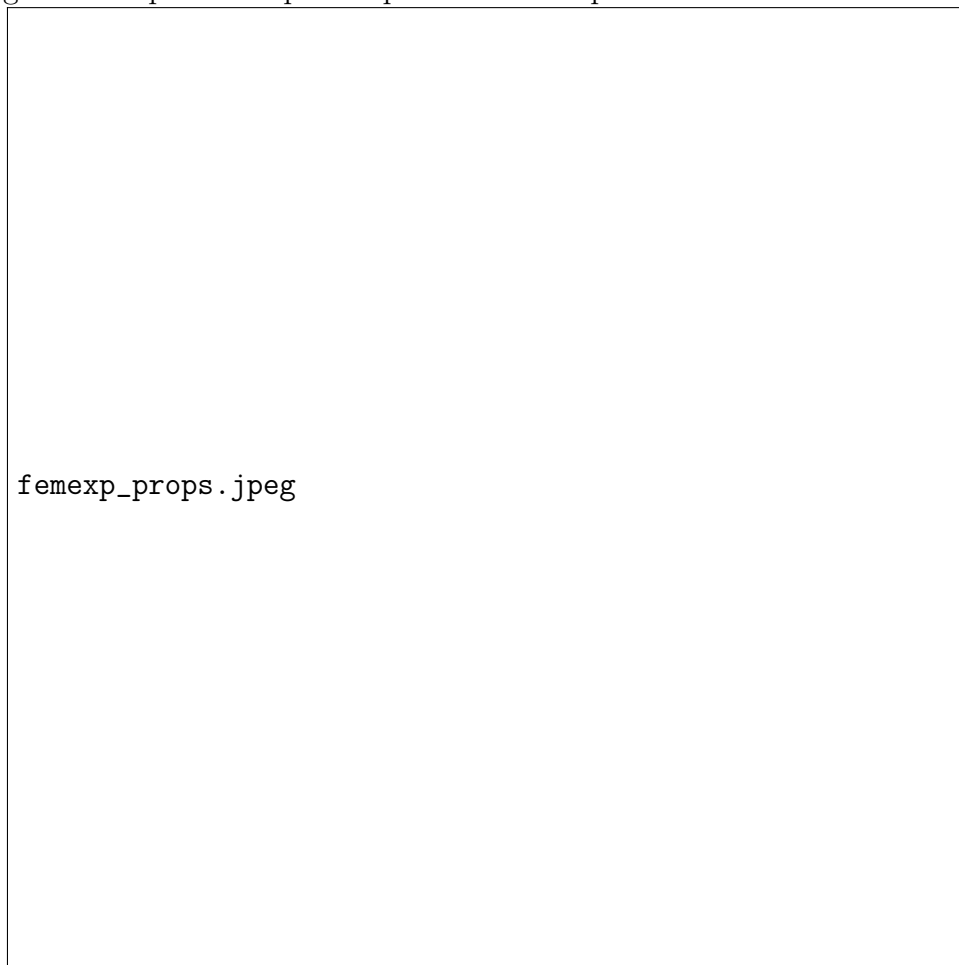
Emphasizing Experience in Elected Office on Candidate Websites

We first employ these regressions to test our experience hypotheses. We conduct difference of means tests to determine if experienced Democratic women reference their electoral background more or less than amateur Democratic women and experienced Democratic men. This analysis focuses on those topic categories that correspond with candidate electoral experience: *Awards & Prestige*, *Electoral Experience*, and *Legislative*.¹¹ Figure 2 displays the top eight topics in candidate website biographies by expected topic proportion. This plot allows us to evaluate the extent to which Democratic women candidates who have previously held elected office talk about their electoral experience in their campaign website biographies. Out of the twenty-seven topics specified by our STM, *Awards & Prestige*, *Electoral Experience*, and *Legislative* are all among the most common topics in the biographies of experienced Democratic women. Figure 2 clearly demonstrates that those topics pertaining to electoral experience are an important facet of self-presentation for experienced Democratic women. To evaluate how the electoral topic prevalence in the bios of experienced Democratic women compare to experienced men and amateur women, we turn next to our difference of means tests.

We expect that experienced Democratic women will have a greater proportion of their campaign biographies dedicated to electoral experience topics than will experienced Democratic men. As previously discussed, women candidates often feel like they must accrue a substantial amount of experience before they can run for office (??). Similarly, women may feel like they must place more emphasis than men on their past political experience during the campaign in order to show they are qualified (??). We also expect that the proportions of experience-related topics will be greater in the campaign biographies of

¹¹Referring to the stems presented in Table 1, both the *Legislative* and *Electoral Experience* topics are clearly indicative of past experience as a legislator or elected official. The *Awards & Prestige* topic is less explicitly tied to electoral experience. We include it nonetheless to ensure a conservative test of our hypotheses.

Figure 2: Expected Topic Proportions for Experienced Democratic Women



This figure depicts a corpus-level visualization of experienced Democratic female biographies. Each proportion is an estimate generated from the structural topic model. The proportions show the amount of the corpus of experienced Democratic female biographies dedicated to each topic category. For example, the most common topic in experienced female Democrats' discourse is the *Legislative* topic (the complete list of topics is presented in Table 1).

experienced Democratic women than in the bios of amateur Democratic women, as they have a political history to draw on.

The results are presented in Table 2. The first row compares the average proportion of the *Awards & Prestige*, *Electoral Experience*, and *Legislative* topics in the bios for experienced versus amateur Democratic women. The second row compares experienced Democratic men and women. The values presented are differences in means with standard errors in parentheses. A negative difference in means indicates that, on average, expe-

Table 2: Difference of Means in Topical Prevalence Across Comparison Groups for Electoral Experience in Campaign Websites

<i>Comparison Group</i>	<i>Topic Category</i>		
	Awards & Prestige	Electoral Experience	Legislative
Democratic Women Experienced vs. Amateur	−0.001 (0.012)	0.041* (0.005)	0.147* (0.010)
Experienced Democrat Women vs. Men	−0.005 (0.008)	0.000 (0.006)	0.014 (0.010)

Values presented are differences in means with standard errors in parenthesis. Negative values indicate that experienced Democratic women had a lower proportion of a given topic in their campaign bios than the comparison group. $p < 0.1$

rienced Democratic women dedicate a smaller proportion of their campaign biographies to a given topic than did the comparison group. Turning to the first row of Table 2, experienced Democratic women were significantly more likely to talk about the *Electoral Experience* and *Legislative* topics in their biographies than amateur Democratic women. This finding aligns with our hypothesis that experienced Democratic women will reference their electoral background more than their amateur counterparts. Comparing experienced Democratic women and men biographies, however, we find no significant difference in the average proportion of these topics. Contrary to our hypothesis, experienced Democratic women do not play up experience more than experienced Democratic men.

Next, we plot the differences in mean topic prevalences for all twenty-seven topics specified by our structural topic model. These full models are displayed in Figure 3. The left panel compares amateur Democratic women to experienced Democratic women. The right panel compares experienced Democratic women to experienced Democratic men. The dashed line denotes the null hypothesis that topic prevalence does not differ significantly across contrasted groups. In Figure 3, topic point estimates falling to the right of the dashed line indicate that experienced women candidates talked more about that

Figure 3: Topic Prevalence for Experienced Democratic Women vs. Comparison Groups

fullmodels.jpeg

Note: The left panel compares amateur Democratic women to experienced Democratic women. The right panel compares experienced Democratic men to experienced Democratic women. The null hypothesis is no difference in topic prevalence between amateur and experienced candidates. Point estimates are generated with 90% confidence intervals.

particular topic. Point estimates falling to the left of the dashed line indicate that the comparison group talked more about that particular topic than did experienced women candidates. Point estimates are ordered by intensity, from most associated with experienced women candidates’ biographies (a positive point estimate) to most associated with the biographies for the comparison group (a negative point estimate). Topic labels included in Figure 3 correspond with those outlined in Table 1.


Several noteworthy patterns emerge in Figure 3. In the left panel, which compares experienced Democratic women and amateur Democratic women, there are nine topics where prevalences diverge. Turning to the right panel of Figure 3, which compares experienced Democratic men and experienced Democratic women, there are only two topics for which candidates diverge.¹² Contrasting these two panels, we find that experienced Democratic women have topical prevalences in their bios more similar to experienced Democratic men than they do to amateur Democratic women. While women candidates may choose to run for Congress using a different decision-making schema than men, once they do decide to run, politically experienced men and women seem to adopt similar self-presentation strategies in their campaigns. This finding calls into question the commonly held perception that most women talk about similar things and present themselves in similar ways when running for office, and that this is different from how men present themselves. Our results suggest that past political experience may be the most salient cleavage in defining how a candidate presents herself to voters.

To illustrate how politically experienced Democrats employ their past electoral experience in their self-presentation, Figures 4 and 5 include several quotes from the biography pages of 2018 congressional candidates. As previously noted, these bios are not simply a rehashing of the candidate’s resume but rather a testament to *why* they should be (re)elected to Congress. In line with expectations, candidates using the *Legislative* and *Electoral Experience* topics communicate their qualification through anecdotes about their

¹²This trend persists for experienced and amateur Democratic men for which there are fifteen topics where prevalences diverge. See the complete model in Figure A.5 of the Appendix.

prior experiences and accomplishments.


Figure 4: Examples of Legislative Topic Use in Candidate Biographies



quotes_leg.png

Without prior experience in elected office to emphasize in their biographies, we next assess the self-presentation strategies of amateur candidates. Looking to the left panel of Figure 3, the biographies of amateur Democratic women are more closely associated with the *Representation* topic than are the biographies of experienced Democratic women. The *Representation* topic is also the most common topic for amateur women and the second

Figure 5: Examples of Electoral Experience Topic Use in Candidate Biographies



quotes_exp.png

most common topic for amateur men. Figure 6 illustrates excerpts from biographies with high proportions of words belonging to the *Representation* topic. In Figure 6, both Ryan Krause (R-TX) and Lisa Ring (D-GA) make appeals using words associated with the types of values that most Americans would like to see in their member of Congress, for instance leadership and integrity. Instead of relying on their experience, these amateurs demonstrate their ability to attain and succeed in elective office through values-driven

Figure 6: Examples of Representation Topic Use in Candidate Biographies



language.

Both experienced and amateur candidates use their biographies to communicate a sense of qualification and aptitude for holding office, however it is clear that they go about it in different ways. Literature on politically experienced candidates finds that these individuals may be more successful in elections because they have fundraising networks,

have previously run a successful campaign, and have greater name recognition (?). From a self-presentation standpoint, these experienced candidates may want to play up their past electoral experience to demonstrate not just their political proficiency, but also to establish a connection to the community. In Figure 5, for instance, Paul Cook (R-CA) tries to make this kind of association, stating that his “neighbors” elected him to the California State Assembly. In doing this, Cook makes a compelling statement, both establishing a relationship with his voters and communicating a political achievement. Amateurs, on the other hand, do not have a demonstrable record of political accomplishments and, therefore, may have no choice but to rely on a rhetorical strategy to communicate their ability to hold office. This first analysis makes clear that emphasizing one’s qualifications and ability to succeed in office is a priority among *all* candidates.

Gender in Candidate Websites

As previously discussed, amateur candidates cannot rely on their electoral background to convey their qualifications for holding office and record of representing constituents. Demonstrating diversity through gender-based appeals may help elevate amateur candidates in lieu of political experience. Unlike amateurs, politically experienced women can employ their political background to win over voters, and therefore do not need to rely as heavily on their gender. While amateur women may be more likely than experienced women to highlight their gender in campaign biographies, we posit that this relationship will be conditioned by party. For Republican women, emphasizing gender may actually serve as a *disadvantageous* strategy because Republican women are often elected not because they are women, but because they are Republicans (?). Therefore, while we generally expect that amateur women will emphasize their gender more than experienced women, we hypothesize that both amateur and experienced Democratic women will talk more about their gender than Republican women.

To evaluate how women candidates employ gender in campaign biographies we focus

on the *Diversity* topic. Referring to Table 1, top FREX stems for this topic include women, advoc, mother, and champion. While this topic deals broadly with diversity, gender is clearly a central component. Figure 7 illustrates excerpts from biographies with high proportions of words belonging to the *Diversity* topic. In Figure 7, candidates in all three examples emphasize their gender and express a commitment to representing women’s issues. Using the *Diversity* topic, we once again conduct difference of means tests to make comparisons across groups and test our hypotheses.

The results are displayed in Table 3, evaluating the prevalence of gender in campaign biographies for Democratic experienced and amateur female candidates. The first column compares the average proportion of the *Diversity* topic in bios for experienced versus amateur Democratic women. The negative difference in means presented indicates that, contrary to our hypothesis, experienced Democratic women dedicate a larger proportion of their campaign biographies to discussions of diversity and gender than did amateur Democratic women. To review, we hypothesized that amateur Democratic women would focus more on gender and diversity since they did not have a history of electoral and political experiences to focus on. However, we actually find that experienced women focused more on gender and diversity in their campaign websites. This may be because these candidates are responding to different electoral contexts (i.e. the type of race the candidate ran in). In 2018, 77% of amateur Democratic women ran in Republican incumbent-held districts and open seats. Conversely, 81% of experienced Democratic women ran in open seats and incumbent-held districts that were controlled by the Democratic Party.¹³

As previous literature suggests, stereotypically “female” issues and traits are associated with the Democratic Party, whereas “male” issues and traits are associated with the Republican Party (???). Therefore, too much emphasis on women’s issues might be seen as a threat to the Republican voters, while taking differing stances on women’s issues might be threatening to many traditionally Democratic, feminist groups (?). As a result,

¹³This number includes incumbents. When excluding incumbents from our measurement of experienced candidates, this number drops to 59%.

Figure 7: Examples of Diversity Topic Use in Candidate Biographies

quotes_diverse.png

amateur Democratic women running in Republican districts may choose to downplay their gender to stand a better chance. Experienced Democratic women overwhelmingly ran in districts friendly to their own party and, therefore, may have been more apt to highlight gender in their self-presentation. By examining an unbalanced sample of candidates, our analysis may be picking up differences in strategy across race type rather than disparities

motivated by candidate political experience or lack thereof.

Table 3: Difference of Means in Topical Prevalence Across Race Type for Gender in Campaign Websites

<i>Topic</i>	<i>Comparing Democratic Females</i>		
	Amateurs vs. Experienced Candidates	Amateurs vs. Experienced in Republican Seats	Amateurs vs. Experienced in Democratic Seats
Diversity & Gender	-0.037* (0.011)	0.003 (0.015)	-0.006 (0.022)

Values presented are differences in means with standard errors in parenthesis. Negative values indicate that the first comparison group had a lower proportion of a given topic in their campaign bios than the second comparison group. $p < 0.1$

To examine this initial finding further we conduct a second series of tests, calculating differences in mean *Diversity* topical prevalence in bios for experienced and amateur Democratic candidates conditional on district type (i.e. Republican incumbent-held district, Democratic-held open seat, etc.). We consider Republican (Democratic) seats to be any district that had a Republican (Democratic) incumbent in 2018 or open seats that were vacated by a Republican (Democrat) before the 2018 primary election. The second and third columns of Table 3 demonstrate that the negative and significant difference in means presented in column one is largely driven by the sample of candidates examined.

The differences between these three tests are likely due to the composition of experienced and inexperienced women in different race contexts. The majority of experienced Democratic women who ran in 2018 emerged in districts safe for their own party whereas amateur female candidates tended to run in Republican-held districts. By breaking these women candidates down by district type we find no statistically significant difference in *Diversity* topic prevalence. Meaning, when comparing experienced and amateur Democratic women who ran in the same type of district, these candidates spent similar amounts of space in their biographies talking about gender. Further, these results indicate that female Democratic candidates tend to customize their self-presentation based on electoral

context. This engages with the broader idea that though Republican women do champion women’s issues (?), they do so in a way that is incongruent with the typical understanding of women’s issues. As a result, Democratic women running in Republican districts are less likely to emphasize gender in a way that risks making them less competitive in their chosen race. One of the benefits of being a candidate with experience is that she knows how to make better campaign decisions, meaning she is more likely to pick a more favorable race type than her inexperienced counterparts. However, our findings show that while amateur candidates might not make the most advantageous decisions about where to emerge, they may be just as strategic in their campaigning, employing their gender only when helpful based on race conditions.

Next, we explicitly test differences between the two parties in how frequently women talk about gender on their campaign websites. In Table 4, we evaluate to what extent Democratic women talk about their gender in campaign bios compared to Republican women running for Congress in 2018. For both experienced and amateur women candidates, Democrats dedicate a larger proportion of their biographies to the *Diversity* topic than do Republican women, aligning with our expectations.

Table 4: Difference of Means in Topical Prevalence Across Comparison Groups for Gender in Campaign Websites

<i>Topic Category</i>	<i>Comparison Group</i>	
	Experience Female Dem. vs. Rep.	Amateur Female Dem. vs. Rep.
Diversity & Gender	0.088* (0.012)	0.037* (0.008)

Values presented are differences in means with standard errors in parenthesis. Negative values indicate that the first comparison group had a lower proportion of a given topic in their campaign bios than the second comparison group. $p < 0.1$

Although we find a difference for the *Diversity* topic, when drawing comparisons for all twenty-seven topics, there are relatively few partisan differences between Democrats

and Republicans of the same experience level. In Figure 8 we present these differences; experienced candidate comparisons are displayed in the left panel and amateur candidate comparisons in the right panel. These plots are truncated to only include those topics where there was a statistically significant difference in topic prevalence between Democratic and Republican women’s biographies. Point estimates falling to the right of the dashed line indicate that Democratic female candidates talked more about that particular topic than did Republican female candidates. Point estimates falling to the left of the dashed line indicate that Republican women talked more about that particular topic than did Democratic women.

A bevy of literature suggests that Democrats and Republicans are systematically distinctive in their campaign self-presentation, running on different types of issues and playing up different personal qualities (???). Examining the kinds of topics for which Republican and Democratic women diverge in Figure 8, Democrats talk more about the *Healthcare* topic — a Democratic party-owned issue — and Republicans talk more about *Taxes/Spending* — a Republican party-owned issue.

Somewhat counter to this literature on gendered differences in campaign presentation, we find few differences in topic prevalences in both panels of Figure 8 than in the left panel Figure 3, which compares candidates from the same party who share the same gender but possess different levels of electoral experience. Recall, when comparing amateur and experienced Democratic women there were nine points of contrast between these comparison groups. In Figure 8, there are only four and five statistically significant topic contrasts for experienced and amateur candidates respectively. For this congressional election, partisanship does not seem to be the divider in campaign presentation style—the divider is past political experience.

Discussion and Conclusion

In our examination of candidate self-presentation in campaign websites, we found more differences within gender groups than across gender groups. In other words, there was great heterogeneity in the presentational styles women used to highlight why they were qualified to hold office in 2018. In particular, political experience divides candidates into two camps with experienced men and women focusing on their legislative experience while amateur men and women are more likely to use values-driven language that emphasizes the aspirational direction in which they would like to move policy. This focus on political experience is far more critical in determining a candidate’s campaign messaging than gender. We find that experienced Democratic women are more stylistically similar to experienced Democratic men than they are to amateur Democratic women. This “experience divide” in presentation is noteworthy, especially since inexperienced candidates are performing well in congressional elections as of late (?). If experienced candidates are emphasizing their political acumen, but then losing the election, it might be worth considering other presentation strategies.

Furthermore, though we found a great deal of heterogeneity among Democratic and Republican women based on levels of experience, we found gendered similarities in comparisons across party. This indicates that there are similarities in the campaign messaging themes that women utilize regardless of party. As the number of women running continues to increase in future elections, it will be interesting to see if women develop more distinctive presentation styles across parties as men do, or if there are certain topics that women—across parties—continue to see as valuable parts of their campaign style.

Overall, our research emphasizes the value of using website data to examine candidate presentation styles. Our work finds that candidate experience is a key factor in determining campaign messaging and something that candidates value as they look for ways to make themselves distinctive by emphasizing the experience that they have. This is particularly true for women. The literature suggests that women only run when they feel

truly qualified to do so, and our findings suggest that this experience is what they choose to emphasize as they work to appeal to voters. In the days and months since the historic 2018 midterm elections, there have been many important conversations about continuing to increase female representation. As more and more women enter the candidate pool, each candidate must distinguish herself from not only the men in the race, but also from other women candidates as well.

Figure 8: Topic Prevalence by Partisanship Conditioned on Electoral Experience for Female Candidates

party_women.jpeg

Note: The left panel compares experienced female Democrats to experienced female Republicans. The right panel compares amateur female Democrats to amateur female Republicans. Null hypothesis is no difference in topic prevalence between candidate comparison groups. Point estimates are generated with 90% confidence intervals. Topics where the difference in prevalence was statistically indistinguishable from zero were omitted for clearer interpretation.

A Appendix

A.1 Pre-Processing & Topic Selection Strategy

Figure A.1: Document, Word, and Token Removal as a function of Threshold



Results were replicated using several different thresholds: 15, 20, and 25. Topics produced and their word associations were nearly identical to those presented in the body of the paper.

We employ a document removal threshold of 30, meaning words that do not appear in at least 30 documents are dropped for computational efficiency. We assume that words appear in such a few number of documents will offer little information for our

analysis. Modifying the document removal threshold threshold to 25, 20, and 15 produced substantively identical results in our subsequent analysis. After removing stop words, punctuation, numbers, and candidate names as well as stemming words to their root, we are left with a corpus of 1,507 documents and 217,536 tokens.

To run our initial STM we set the model initialization to **"Spectral,"** which uses the connection of LDA with non-negative matrix factorization that provides theoretical guarantees that the optimal parameters will be recovered. We set the number of topics, K , to zero, which automatically uses an algorithm developed by ? to select the number of topics. ? stress that this approach does not select the “true” number of topics but rather is a useful place to start. This initial model produced a total of 48 topics.

After finding an initial number of topics, we ran several STMs, varying the number of topics across a sequence of values from 10 to 60 using the **searchK** function in the **stm** package. The results are presented in Figure A.2. Four metrics are provided to assess the model quality: held-out likelihood, residual dispersion, topic semantic coherence and the approximation to the lower-bound of the marginal likelihood. The goal is maximize the held-out likelihood, semantic coherence, and lower-bound while minimizing the dispersion of residuals. Referencing Figure A.2, we ran two more **searchK** functions narrowing our sequence of topics each time based on metric performance. Our final run included a sequence of K topics from 20 to 40. The results are presented in Figure A.3. We settled on specifying our model with 27 topics. Modifying the number of topics slightly, increasing and decreasing K by one produced results nearly identical to those presented in the body of the paper.

To determine the strength of our modeled topics we produced a plot of topic quality, which is displayed in Figure A.4. Topic quality is evaluated using semantic coherence — how often words within a topic co-occur — and exclusivity — the uniqueness of words to each topic — which are displayed on the x-axis and y-axis respectively. Using this approach is accepted as a reasonable surrogate for human judgment on the quality of

topics (?). The highest quality topics fall in the top right corner of Figure A.4. Based on the semantic coherence and exclusivity metrics, around six topic produced by our model, including *Nature/Land*, *Immigration*, *Agriculture*, and *Partisan Issues* could be considered “lower-quality.” However, reviewing topic prevalences in Table 1, these lower quality topics are also those that occur less often in the text, which lessens concern that our interpretation of results is not meaningful.

A.2 Covariates, Sample & Results

The following section includes results that were discussed but not presented in the body of the paper. Topics presented in Table A.3 are produced with the exact same model as Table 1, the only difference is the second column where highest frequency word summaries for topics are presented in lieu of FREX word topic summaries.

Table A.1: Structural Topic Model Covariate Summary

Factor Level	Gender	Party	Electoral Experience
0	Male	Republican	Experienced Candidate or Incumbent
1	Male	Republican	Political Amateur
2	Female	Republican	Experienced Candidate or Incumbent
3	Female	Republican	Political Amateur
4	Male	Democrat	Experienced Candidate or Incumbent
5	Male	Democrat	Political Amateur
6	Female	Democrat	Experienced Candidate or Incumbent
7	Female	Democrat	Political Amateur

Table A.2: Percent of Candidates Across Each Race Type, By Gender

Gender	Republican Incumbent	Democrat Incumbent	Open Seat
Male	47.58% 275	26.99% 156	25.43% 147
Female	44.14% 162	25.89% 95	29.97% 110

Table A.3: Topics as a Function of Gender, Party, and Political Experience using Highest Frequency Words

Description	Stems	%
Family / Education	school, work, year, colleg, famili, graduat, univers, high, rais, children	0.08
Awards / Prestige	univers, serv, member, board, school, associ, award, degre, presid, state	0.07
Representation	peopl, work, need, repres, believ, communiti, govern, district, want	0.07
American Dream	work, famili, busi, american, hard, small, job, dream, help, valu, start, communiti, opportun, grow, compani, generat, live, parent, creat, rais	0.06
Business / Industry	busi, develop, manag, experi, communiti, local, year, career, industri, econom, includ, leadership, compani, servic, project, director, public	0.05
Legislative	committe, hous, serv, legisl, member, congress, repres, also, congressman	0.05
Military	serv, servic, militari, armi, forc, unit, air, veteran, offic, state	0.05
Tax / Spending	tax, govern, conserv, busi, state, small, spend, budget, fight, right, nation	0.05
Think-Say-Do	can, get, like, time, just, make, peopl, know, one, live, need, thing, want, life, back, take, politician, come, even, say	0.05
Bipartisanship	job, work, help, congress, creat, veteran, improv, new, economi, make, secur, ensur, get, local, support, communiti, feder	0.04
Elect. Experience	district, elect, repres, serv, counti, citi, congression, state, year, member	0.04
Healthcare	health, care, educ, afford, healthcar, protect, access, right, qualiti, system	0.04
Activism	right, progress, fight, polit, corpor, chang, money, work, campaign, peopl, interest, power, worker, democrat, climat	0.03
Blue Collar	work, year, compani, manag, busi, engin, program, start, industri, time, experi, high, move, financi, mani, sever	0.03
Campaign Phrases	congress, run, district, time, america, home, now, meet, futur, congression	0.03
Diversity	communiti, educ, work, women, school, organ, advoc, program, student, public, help, children, support, first, state, democrat, polici, campaign, immigr	0.03
Medicine / Doctor	medic, health, care, doctor, research, univers, hospit, medicin, cancer, nurs, practic, patient	0.03
Legal / Lawyer	attorney, crimin, state, justic, offic, polic, counti, legal, crime, practic, serv	0.03
President / Party	republican, democrat, parti, polit, trump, presid, candid, campaign, american	0.03
Religion	year, life, state, serv, church, children, america, faith, unit, also, god	0.03
Support / Donate	vote, read, campaign, district, elect, donat, support, candid, pleas, repres	0.03
Agriculture	farm, counti, agricultur, volunt, farmer, cross, local, western, river, state, march, central, organ, valley, island, north, rural, involv, blue, fish	0.02
Coaching	play, team, coach, high, scout, sport, success, boy, tough, help	0.02
Terrorism	secur, nation, american, america, defens, presid, foreign, polici, war, world	0.02
Nature / Land	servic, face, resourc, water, open, parti, land, third, natur, generat, sourc, fourth, valley, light, enjoy, protect, bold, behalf, rang	0.01
Immigration	issu, paid, support, import, border, post, name, point, focus, top, congress, status, activ, divid	0.01
Partisan Issues	state, can, america, year, peopl, school, gun, unit, program, cost, need, use, provid, social, million, educ, must, tax, american, make	0.01

Figure A.2: Diagnostic Values for STM Model (Sequence of Topics 10-60)



Model value K varies across the sequence: 10, 20, 30, 40, 50, 60. Model specified with factor covariate from Table A.1 and successfully converged

Figure A.3: Diagnostic Values for STM Model (Sequence of Topics 20-40)



Model value K varies across the sequence of integers 20 through 30. Model specified with factor covariate from Table A.1 and successfully converged.



Figure A.4: Partisan-Gender-Experience Model Topic Semantic Coherence and Exclusivity

Note: Semantic coherence, on the x-axis, refers to how often words within a topic co-occur. Exclusivity, on the y-axis, refer to the uniqueness of words to each topic. Topics falling in the upper right corner have the highest “quality” maximizing both exclusivity and semantic coherence.

Figure A.5: Topic Prevalence by Partisanship Conditioned on Electoral Experience for Male Candidates



Note: The panel compares amateur Democratic men to experienced Democratic men. Null hypothesis is no difference in topic prevalence between candidate comparison groups. Point estimates are generated with 90% confidence intervals. Topics where the difference in prevalence was statistically indistinguishable from zero were omitted for clearer interpretation.

Figure A.6: Topic Prevalence for Experienced Republican Women vs. Comparison Groups



Note: The left panel compares amateur Republican women to experienced Republican women. The right panel compares experienced Republican men to experienced Republican women. The null hypothesis is no difference in topic prevalence between amateur and experienced candidates. Point estimates are generated with 90% confidence intervals.

Figure A.7: Topic Prevalence by Partisanship Conditioned on Electoral Experience for Male Candidates



Note: The left panel compares experienced male Democrats to experienced male Republicans. The right panel compares amateur male Democrats to amateur male Republicans. Null hypothesis is no difference in topic prevalence between candidate comparison groups. Point estimates are generated with 90% confidence intervals. Topics where the difference in prevalence was statistically indistinguishable from zero were omitted for clearer interpretation.