DATA FOR PROGRESS
OCTOBER 2018

SOMETHING TO RUN FOR: CANDIDATE EMERGENCE IN THE PROGRESSIVE MOVEMENT

MEREDITH CONROY AND JON GREEN

Since President Trump's inauguration, tens of thousands of individuals with little or no previous political experience have expressed interest in running for office at all levels of government, including as many as 40,000 women (Alter 2018). This interest among women has translated into a surge in women-led candidacies; in 2018, there was an 88 percent increase in women running for Congress relative to the previous midterm year (Center for American Women and Politics 2018). There has also been a surge in LGBTQ candidates (Tumulty 2018), scientists (Yong 2018), and young people (Schneider 2018) running for office. Although there has been a lot of speculation about why individuals from historically underrepresented groups have decided to enter the political fray to such a degree, there has yet to be an effort to quantitatively verify some of these theories. In this report, we partner with Run for Something, a non-profit with the goal of recruiting young and diverse progressive candidates to run for political office, to explain the emergence of new candidates in the progressive movement, and to identify the sorts of traits and motivations that are indicative of translating interest into a real campaign, at any level of government.

EXECUTIVE SUMMARY

Run for Something provided Data for Progress with over 13,500 de-identified[1] intake forms from their website collected between April of 2017 and July of 2018. The intake form was accessible on the Run for Something's website, where individuals could visit and submit a form notifying the organization of their interest in potentially becoming a candidate. Forms asked respondents to respond to a wide variety of fields, including birth year, race, gender, ZIP code, the reasons why prospective candidates are interested in running for office, and tags Run for Something entered in association with their continued involvement with the organization. Data for Progress analyzed these data to investigate which prospective candidates were most likely to emerge as actual candidates, and how they articulated their interest in running.

Key Findings:

- **REPRESENTATION MATTERS:** Women and people of color were more likely to express concerns regarding descriptive representation. White men were more likely to discuss partisan, establishment, or other general political dynamics that were not tied to their race or gender.
- **CLOSING THE GAP:** Conditional on filling out a form, prospective candidates who were identifiable as non-white were more likely to run for office than their white counterparts.
- **AGE MATTERS MORE FOR WOMEN THAN MEN:** Consistent with prior academic findings (Thomas 2002; Fulton et al. 2006), young men were slightly more likely to run than young women, but the likelihood of running increased more for women as age increased than for men.
- **RURAL CANDIDATES KNOW THEY HAVE A STEEP HILL TO CLIMB:** Rural prospective candidates were particularly likely to mention the political dynamics in their communities – especially how conservative they generally are – but those who did were also relatively more likely to run for office than those who did not.
- **CANDIDATES TALK ABOUT ISSUES:** Very few prospective candidates articulated their interest in running for office in terms of Trump, and those who did mention him (or the broader political climate) were less likely to run. Candidates were more likely to articulate their interest in specific terms, mentioning their communities' needs, issues they care about (frequently education and health care), and the political dynamics of the districts in which they were considering running.
- **WHAT PROSPECTIVE CANDIDATES SAY MATTERS MORE THAN WHO THEY ARE:** Modeling which prospective candidates became actual candidates shows that the information contained in the statements of interest vastly outperforms the demographic information in the rest of the forms.

THE DATA

Data for Progress re-coded a handful of fields from the intake form in order to prepare the data for analysis. Gender responses were recoded to a three-category variable: "f" if the respondent clearly identified as female, "m" if they clearly identified as male, and "o" if the respondent either affirmatively identified as non-binary or if they did not indicate a gender at all. Similarly, race was recoded to "w" if the respondent identified as white, "n" if the respondent identified with a race other than white, and "u" if the respondent did not identify with any racial group. Respondents were categorized as urban, suburban, or rural (USR) based on their ZIP code, using the Census Bureau's classification. We also constructed respondent age based on their birth year.

Since we are particularly interested in candidate emergence – as in, moving from mere interest in running for office to actually doing so – we identify candidates based on tags, assigned by Run for Something, that are indicative of such behavior. These include "filed to run", "applied for endorsement", "endorsed candidate," and a handful of related tags provided by the organization.

> "I am sick and tired of seeing our country continuously put in a place that sets my generation up for failure. We have to recover from that now while we still can.
>
> 27 F

DESCRIPTIVE ANALYSIS

Overall, there are 13,697 intake forms in the data. For analysis, we subset down to respondents who were at least 18 years old and wrote at least four words in their statement of interest in running for office (10,384 cases). We briefly include descriptive tables of the proportion of these respondents who fall into each gender, racial, and geographic category below. As the tables show, these respondents are somewhat more likely to be male than female (though a third of respondents do not have an identifiable gender – the vast majority of these unidentifiable cases left the item completely blank), much more likely to be white than non-white (though a majority of respondents do not have an identifiable race), and three quarters of them reside in suburban ZIP codes.

As Tables 1 and 2 show, women and people of color account for a greater share of candidates than they do of respondents – in part because candidates were much more likely to have an identifiable race and gender. As shown in Table 4, in line with Run for Something's stated goal of recruiting a diverse slate of candidates, thirty percent of women of color who submitted an intake form wound up running for office, while twenty percent of white men did so. While there were slight differences in the geographic distribution of candidates relative to all respondents – Table 3 shows that candidates are more likely to be from rural areas and less likely to be from urban areas – the clear majority of both respondents and candidates are suburban.

TABLE 1 | RUN FOR SOMETHING GENDER DISTRIBUTION

	FEMALE	MALE	UNKNOWN
All Respondents	28%	38%	34%
Candidates	37%	47%	15%

TABLE 2 | RUN FOR SOMETHING RACIAL DISTRIBUTION

	NON-WHITE	WHITE	UNKNOWN
All Respondents	10%	32%	58%
Candidates	16%	40%	43%

TABLE 3 | RUN FOR SOMETHING GEOGRAPHIC DISTRIBUTION

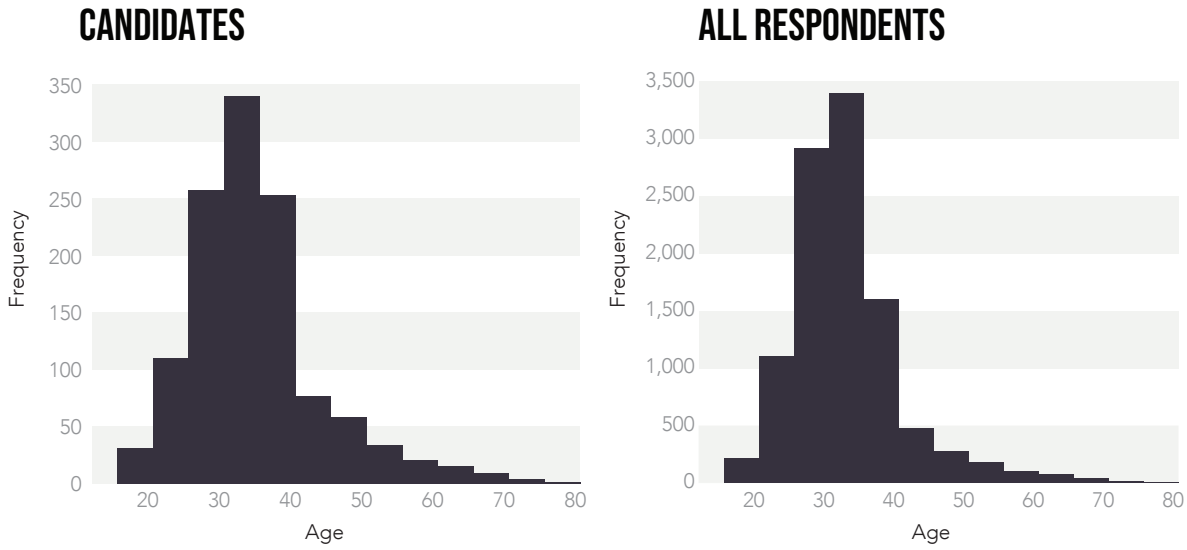
	RURAL	SUBURBAN	URBAN	UNKNOWN
All Respondents	9%	75%	15%	1%
Candidates	13%	74%	12%	1%

TABLE 4 | PROPORTION RUNNING BY RACE AND GENDER

RACE/GENDER	PROPORTION RUNNING
Women of Color	30%
Men of Color	28%
White Women	24%
White Men	20%

Note: Data from respondents who had identifiable race and gender.

We include the distribution of respondents' ages below. The median respondent is 32 years old and three quarters of respondents are 36 years old or younger, consistent with Run for Something's stated goal of generating interest in running for office among young people. The full range of respondent ages runs from 18 to 77.



The age distribution of candidates is only slightly different than the age distribution of overall respondents, with a median age of 34 and a third quartile age of 38. The age range of candidates runs from 18 to 77.

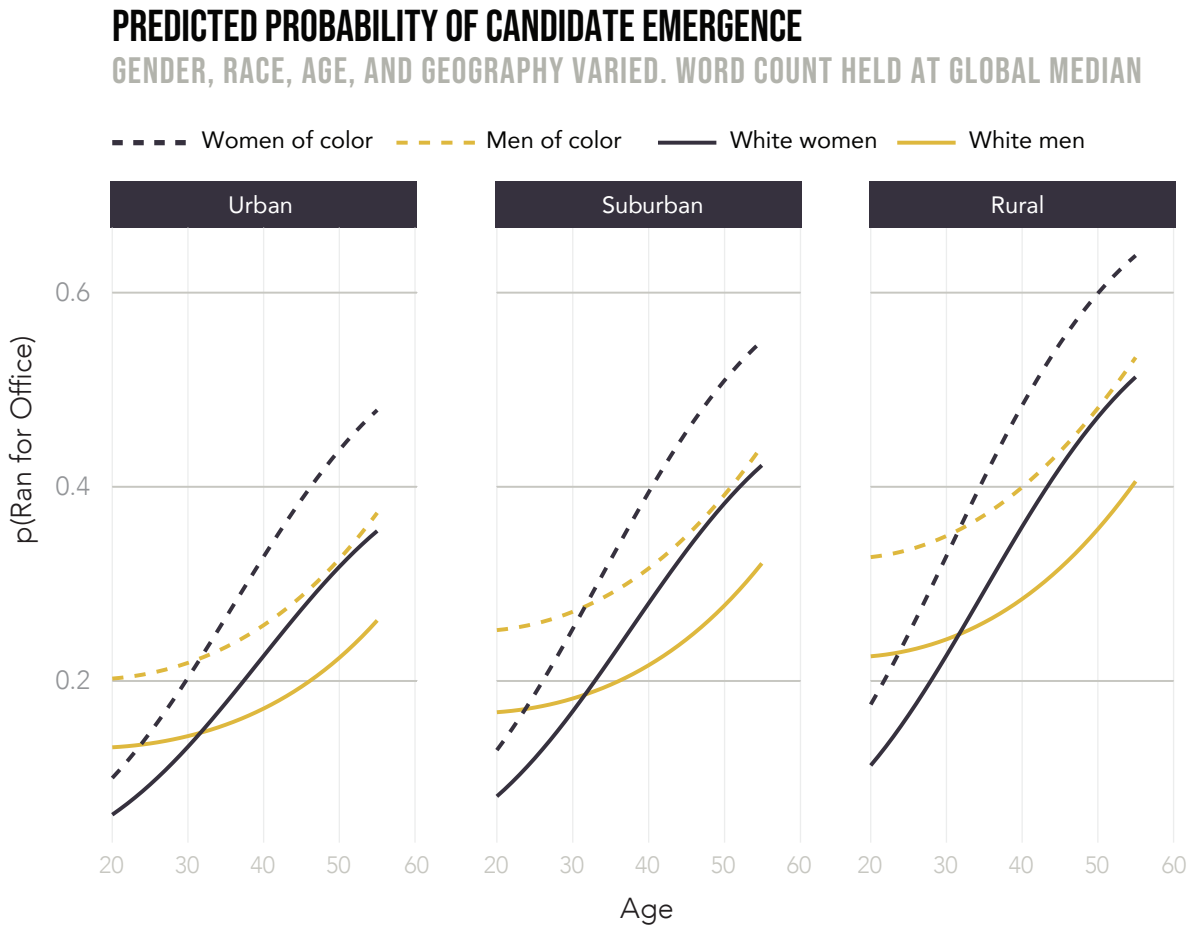
MODELING CANDIDATE EMERGENCE

Roughly ten percent of the respondents in the dataset contain a tag indicating that they ran for office. In order to model candidate emergence conditional on the other factors in the dataset, we specify a rare events logistic regression with independent variables gender, race, USR, age, age squared, the log of the word count in the respondents' statements of interest, and interaction terms between both age variables and gender. The non-linear age term and age-based interactions are included to account for the possibility that age effects are not uniform, either across different groups which may experience different societal expectations based on their age or across age itself.

Selected results from this model are plotted below, where we predict the probability of running for office for each race/gender/geography/age combination for respondents between the ages of 20 and 55, holding word count constant at its global median (46 words). Geographic region is broken out by facet, gender is represented by line color (black for women, gray for men), and race is represented by line type (solid for white, dashed for non-white).

The plot reflects many of the significant relationships uncovered by the model. There are main effects for race, with non-white respondents having a higher marginal probability of running for office than white respondents. Additionally, there are significant interaction effects between gender and both age variables, indicated by the steeper slope and sharper concavity for female respondents' predicted probabilities of running for office as age varies. Put another way, while men at the low end of the age range are predicted to be slightly more likely to run than women of the same age, the probability of running increases with age more rapidly for women than it does for men. We also note a slightly higher overall probability of candidate emergence in rural ZIP codes relative to urban ZIP codes, perhaps indicating fewer progressive candidates emerging in these areas through the conventional candidate pipeline such that there is less preventing candidates from converting interest into action in these regions.

> I want to usher in a new age of political awareness directed to the youth and help historically disenfranchised communities.
>
> 20 M



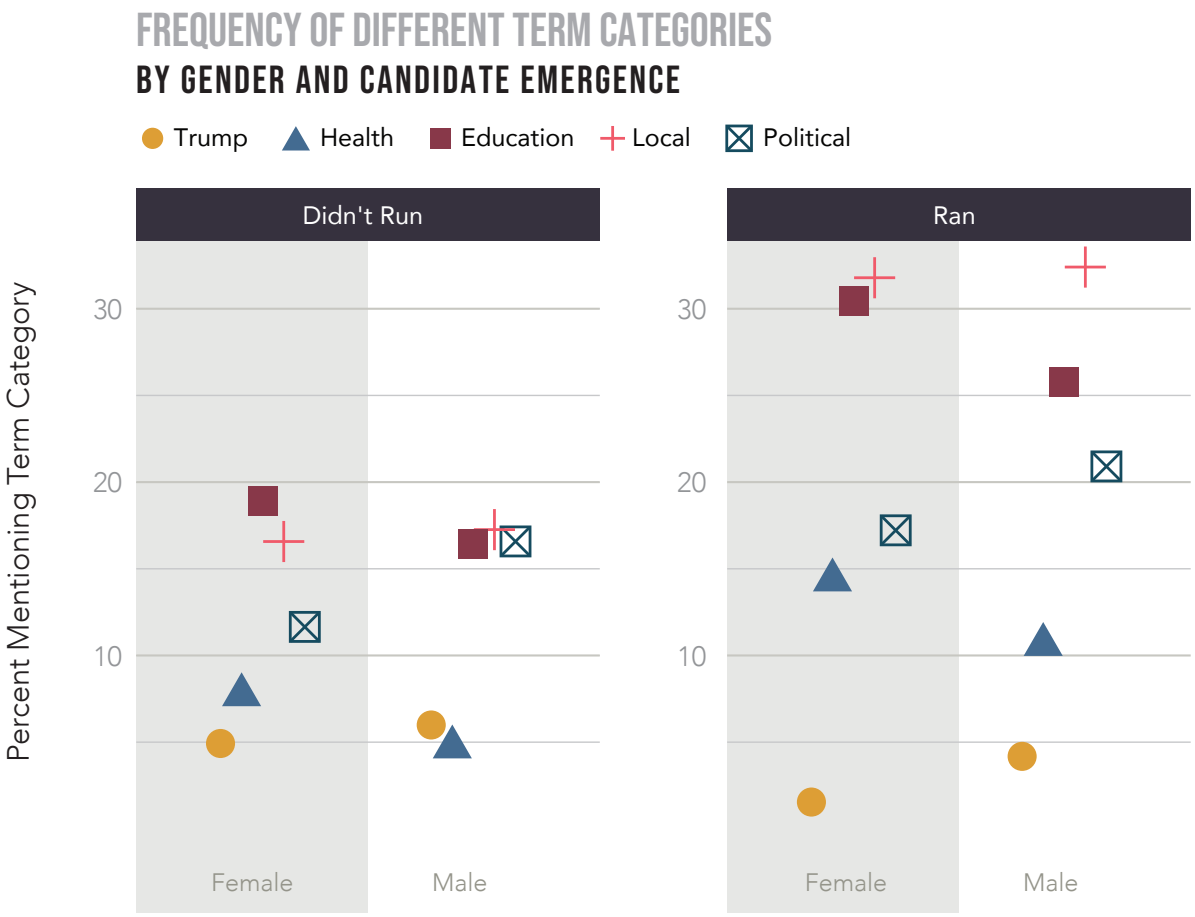
It's important to note that while many of these relationships are statistically significant, on balance this demographics-only model is not very good at predicting candidate emergence. In order to get better leverage on which candidates eventually do run for office, we incorporate information from the statements of interest.

PRELIMINARY TEXT ANALYSIS

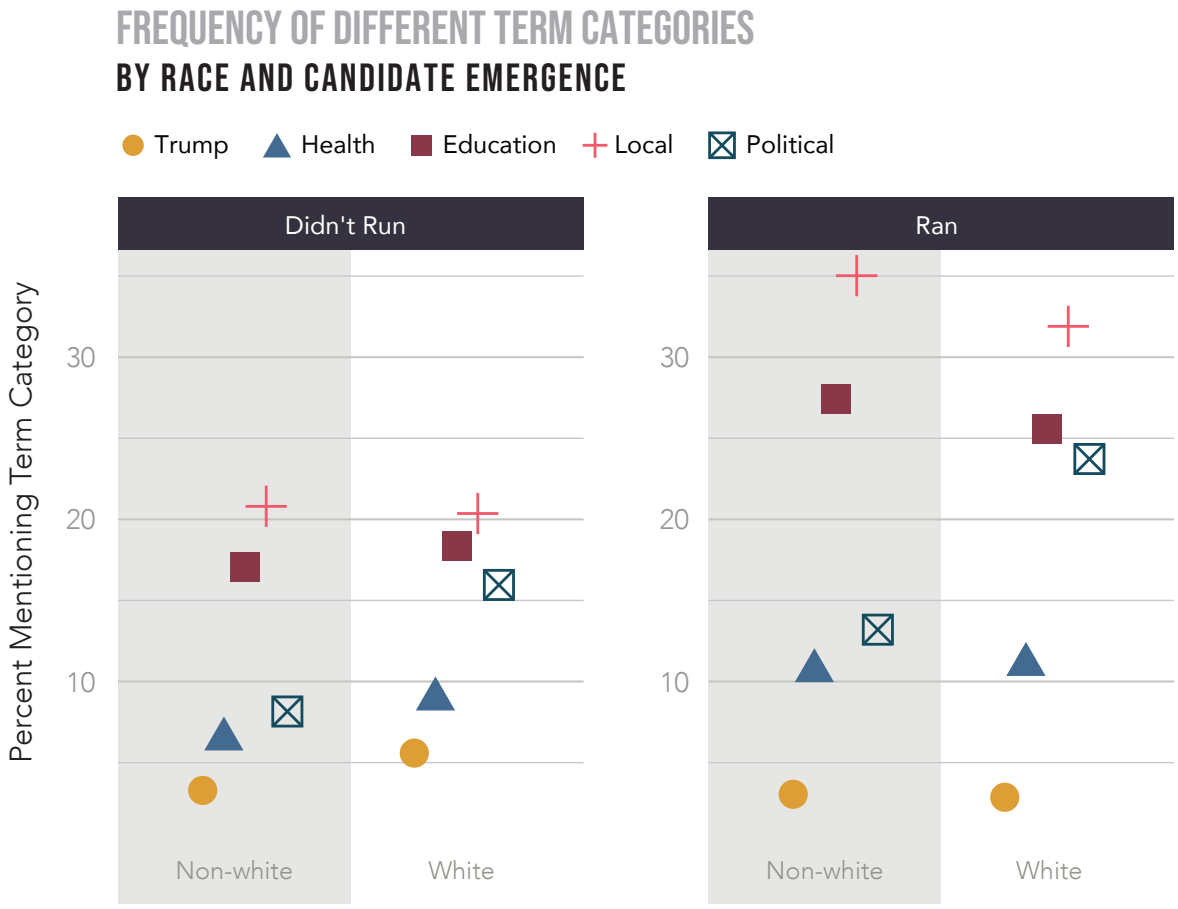
We begin by simply specifying a set of term categories – groups of (in some cases stemmed) keywords we are interested in – to see if different types of respondents are more likely to mention them. The term categories are:

- Trump: trump, fascist, cheeto, darkness
- Health: health, insurance, medical
- Education: educat, school
- Local: local, municipal, county, city
- Political: liberal, democrat, dems, republican, conservative, gop

We first plot the prevalence of these term categories by gender and candidate emergence. As the plot shows, both male and female respondents who actually ran for office were more likely to mention health, education, local, and political terms. While explicit mentions of Trump were relatively rare, female candidates were particularly unlikely to use words that fall into that term category. Female candidates were also about five percentage points more likely to mention education than male candidates, while male candidates were slightly more likely to mention political terms than their female counterparts.

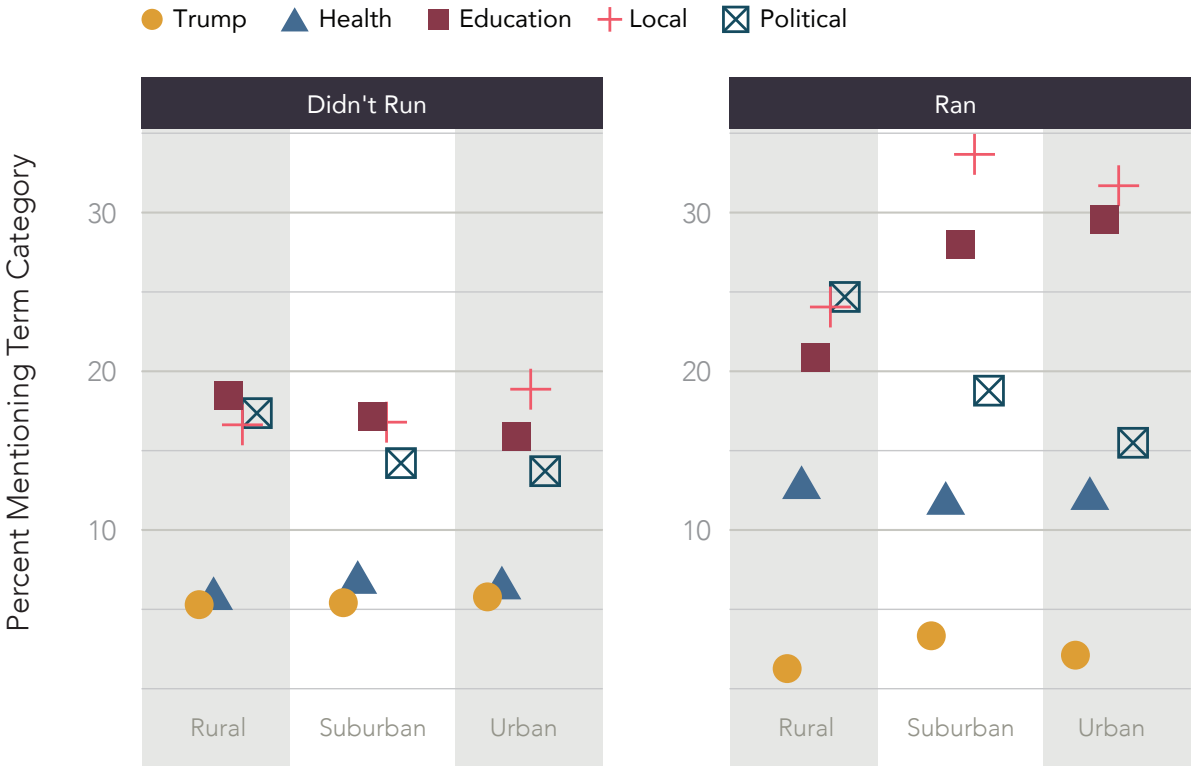


Repeating this method by racial identification, we again see that candidates are more likely to mention local, political, education, and (to a lesser extent) health-related terms than non-candidates. The only obvious racial divide in term categories mentioned is that white respondents – both candidates and non-candidates – are much more likely to mention political terms than non-whites. It is important to note here that this plot is only based on respondents who had an identifiable race, which is only about half of the full dataset.



Finally, we repeat this method by geographic context, finding few differences in term categories used by non-candidates across geographic context. However, we do find that rural candidates are somewhat more likely to mention political terms in their reasons for running, while urban and suburban candidates are more likely to mention local and education-related terms.

FREQUENCY OF DIFFERENT TERM CATEGORIES
BY GEOGRAPHIC CONTEXT AND CANDIDATE EMERGENCE



However, these crosstabs can only convey so much information regarding how different respondents articulated their interest in running for office. In the next section, we incorporate as much information as possible by using recently-developed modeling techniques in order to construct a more complete depiction of what has driven progressive interest in and conversion to running for office over the past year or so.

> I am a woman who is independent. I'm a Muslim revert. And I don't take no for an answer.
>
> 35 F

TOPIC MODELING

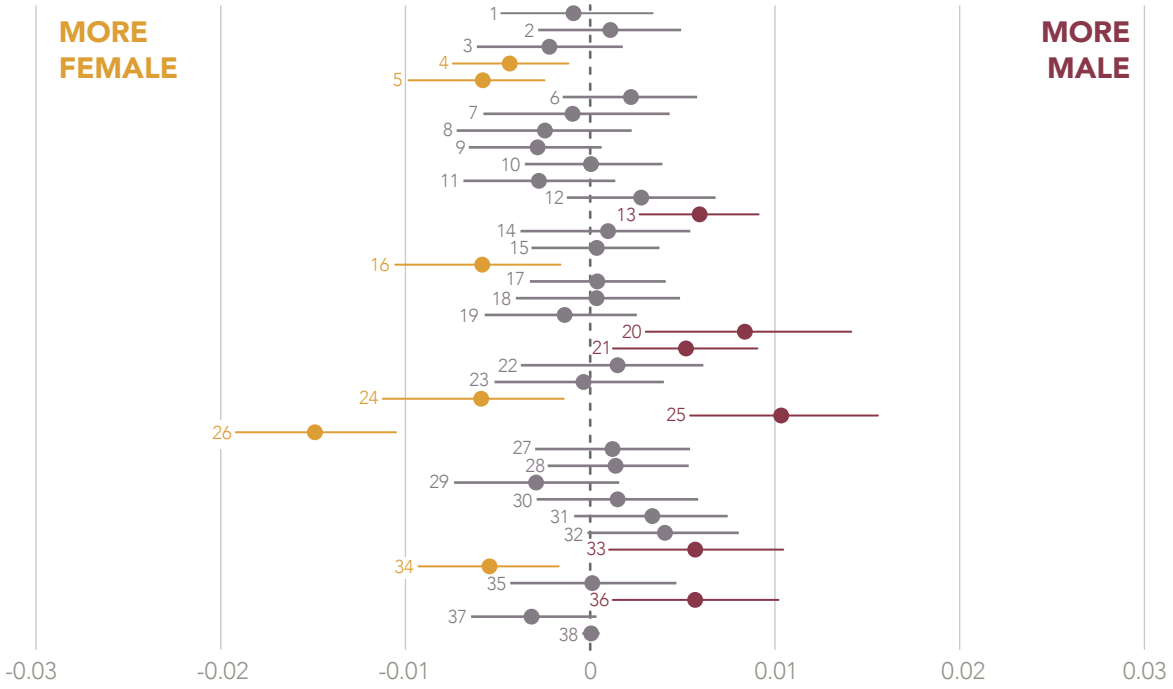
We begin by specifying a structural topic model (STM), which identifies the different things prospective candidates talked about in their statements of interest. Unlike previous topic modeling approaches, the STM allows us to incorporate the metadata from the intake forms – specifically, respondents' race, gender, age, geographic region, verbosity, and candidate status – to identify both topic prevalence and the topics themselves.

While there is no "true" answer to the question of how many topics is the right number of topics to specify in a model such as this one, there are a few metrics we can use to help us land on a model that appropriately balances fit and parsimony. Using these metrics, we have landed on 38 as being the appropriate number of topics to fit a STM to this dataset.

Below, we plot the relative influence of selected covariate levels on topic prevalence. In other words, these plots show how different respondent characteristics are associated with the relative likelihood of that respondent mentioning particular topics. Some topics were mentioned more frequently by female respondents, some by male respondents; some by urban respondents, some by rural respondents; and so on.

We first plot the relative topic prevalence by male vs. female. While the effects of gender on topic prevalence are for the most part substantively small or negligible, there are a few notable exceptions. In particular, topics 13, 20, 25, 33, and 36 are clearly associated with male respondents, and topics 5, 16, 24, 26, and 34 are associated with female respondents:

EFFECT OF FEMALE VS. MALE



Of course, "Topic 26" is, in a vacuum, not very informative. What does it mean for these topics to be associated with gender covariates in substantive terms? Below, we show the words most closely associated with each of the topics listed above as being significantly associated with a gender covariate.[2] First, for men (topics 20, 25, 33, and 36), then for women (topics 5, 16, 24, 26, and 34).

```
## TOPIC 20 TOP WORDS
## state, republican, congress, district, senate, hous, incumb, congress, unoppos, pass, governor, wisconsin, gerrymand
## TOPIC 25 TOP WORDS
## democrat, progress, state, counti, parti, republican, town, bench, committe, win
## TOPIC 33 TOP WORDS
## administr, agenda, angri, backward, countri, current, trump, move, forward, progress
## TOPIC 36 TOP WORDS
## america, stand, presid, divis, divid, believ, america, need, time, must, principl
```

Generally speaking, the topics associated with male respondents tend to focus on political dynamics, either partisan words (republican, democrat), words alluding to pragmatic political concerns (gerrymander, unopposed, incumbent, bench), or words denoting frustration with the current political dynamic (trump, divide, current, backward). As one male respondent whose statement of interest was particularly representative of topic 25 wrote (county location anonymized), "I am running to bring a balance back to [my county] government. It has been 20+ years since a Democrat has won a partisan election in this county. The 2016 for the presidential race show that there are sufficient Democrats in the county and if we GOTV, we can win. If we can turn [my county] purple, then we can turn Texas purple. Change begins with me, and I am proud to be part of the Blue Wave."

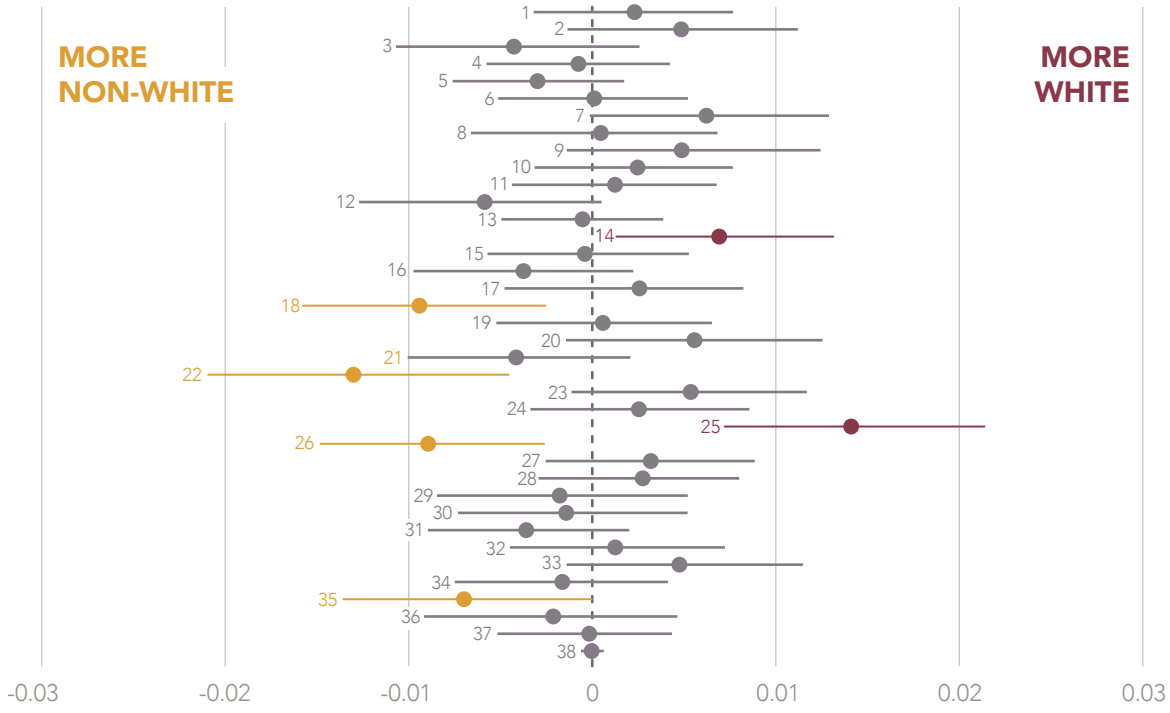
Next, we repeat this process for female respondents, finding that the topics they are more likely to mention tend to relate to district and issue-specific representation, with topics denoted by words including health, education, immigration, and lgbt. In addition, topic 34 includes the words "women" and "underrepresented," potentially indicating that female respondents may have been more likely to articulate their interest in running in terms of descriptive representation.

```
## TOPIC 5 TOP WORDS
## repres, need, district, state, someon, candid, constitu, represent, someon, touch, desper
## TOPIC 16 TOP WORDS
## want, help, differ, make, communiti, abl, speak, voiceless, show
## TOPIC 24 TOP WORDS
## polici, social, health, econom, issu, public, work, mental, oklahoma, equiti
## TOPIC 26 TOP WORDS
## woman, immigr, mother, man, gay, color, ask, black, queer, answer
## TOPIC 34 TOP WORDS
## voic, progress, valu, women, fight, want, heard, platform, lgbt, underrepres
```

As one female respondent whose statement of interest was particularly representative of topic 34 wrote, "With my research background in sociology and health, I am interested in promoting positive social and policy changes to reduce inequalities and promote equity and protections for marginalized groups."

Next, we plot the relative topic prevalence by white vs. non-white. However, we note that the majority of respondents did not have an identifiable race, and are not represented in this plot. Again, we find that most topics are not significantly associated with the respondent's race, but that there are a few notable exceptions. In particular, topics 18, 22, and 26 are more prevalent among non-white respondents, while topics 14 and 25 are significantly more likely to be mentioned by white respondents:

EFFECT OF NON-WHITE VS. WHITE



I've always wanted to but it's never been the right time. There is no right time, come to find out. Now is the time.

36 F

Below, we show the words associated with the topics more frequently mentioned by white respondents. It is worth noting that topic 25 was also more frequently mentioned by male respondents, and concerns partisan dynamics in the respondent's district. The other topic associated with white respondents, topic 14, includes words such as sick, tired, politician, and bystander – indicating a general sense of frustration with the status quo.

```
##  TOPIC 14 TOP WORDS
##  tire, see, just, politician, get, sit, thing, sidelin, sick, wait, complain
##  TOPIC 25 TOP WORDS
##  democrat, progress, state, counti, parti, republican, town, bench, committe, win
```

As one white respondent whose statement of interest was particularly representative of topic 14 wrote, "I'm tired of the status quo. I'm tired of watching as entrenched big money buys more laws that prevent the dismantlement of this archaic cronyism holding the world back. I'm tired of sitting back and expecting someone else to solve the world's problems." In fact, three of the top five documents associated with topic 14 begin with the words "I'm tired."

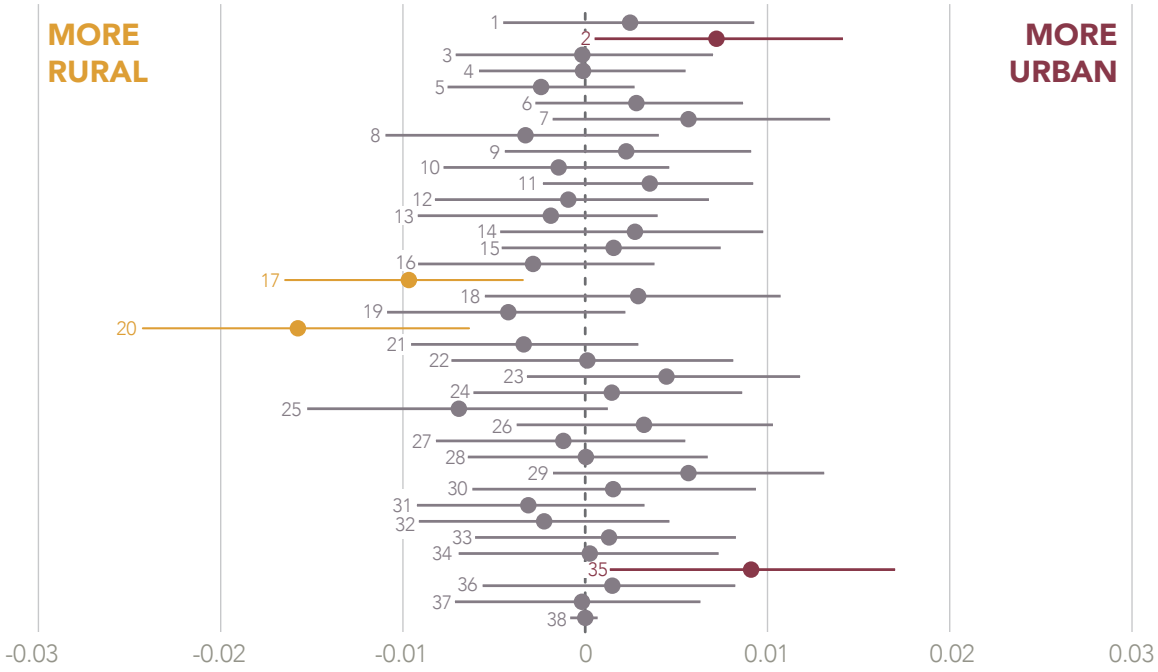
Repeating this method for topics associated with non-white respondents (18, 22, and 26), we find that the topics that are particularly likely to be mentioned by non-white respondents are associated with community, service, and leadership. We also note that topic 26, which contains a number of identity-based words including immigrant, black, gay, and Muslim, was also more prevalent among women as shown above.

```
##  TOPIC 18 TOP WORDS
##  communiti, serv, public, servic, work, advoc, organ, dedic, veteran, sector, servant
##  TOPIC 22 TOP WORDS
##  need, communiti, one, leader, leadership, believ, strong, matter, margin, engag, civic
##  TOPIC 26 TOP WORDS
##  woman, immigr, mother, man, gay, color, ask, muslim, queer, answer
```

As one non-white respondent whose statement of interest reflected topic 22 wrote, "My community needs bold leaders willing to go to bat for them. I want to be one of those leaders."

Next, we plot the relative topic prevalence by urban vs. rural. It is important to note here that the majority of respondents were in suburban ZIP codes, and are not represented in this plot. Here, we find that topic 20 is significantly more prevalent among rural respondents, while topic 35 is more prevalent among urban respondents.

EFFECT OF RURAL VS. URBAN



We plot the words and top documents associated with each of these topics together. As these topics indicate, rural respondents (topics 17 and 20) are more likely to be concerned with the political dynamics in their specific districts or states – perhaps due to the increasing association between rurality and conservative political success – while local issues specific to cities are more likely to be mentioned by urban respondents (topic 35).

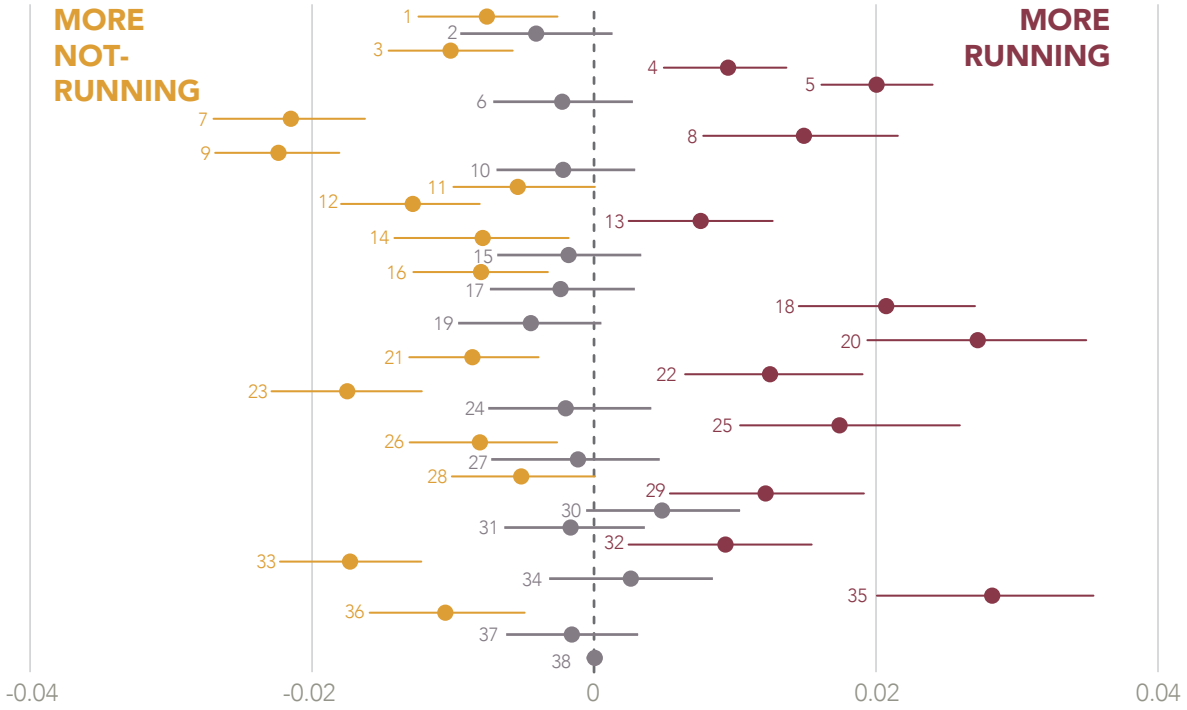
```
## TOPIC 17 TOP WORDS
## area, live, rural, wonder, avail, nativ, san, bay, massiv
## TOPIC 20 TOP WORDS
## state, republican, congress, district, senat, hous, incumb, unoppos, pass, governor, wisconsin, gerrymand
## TOPIC 35 TOP WORDS
## citi, council, resid, neighborhood, mayor, run, develop, ward, baltimor
```

As one rural respondent whose statement of interest was representative of topic 20 wrote, "My state is in serious trouble, we have a 500 million dollar deficit in 2017 and all the Republican legislature wants to do is cut the budget even more, get rid of income tax and switch [entirely] over to a increased sales tax."

An urban respondent, whose statement of interest was representative of topic 35 wrote, "I would like to run for City Council. Our city is growing leaps and bounds, but more focus on expensive houses to increase property taxes and little to no affordable houses or rentals. We also have no one to hold landlords accountable on the maintenance of the rental properties."

Finally, we plot relative topic prevalence by running vs. not running. Unlike with the demographic covariates, here we find a large number of significant relationships between topic prevalence and candidate emergence that point in opposing directions from topic to topic. Specifically, respondents who wrote more words in topics 4, 8, 13, 18, 20, 22, 25, 29, 32, and 35 were significantly more likely to have run for office; respondents who wrote more words in topics 7, 9, 11, 23, 33, and 36 were significantly less likely to do so.

EFFECT OF NOT RUNNING VS. RUNNING



Some of these topics were also associated with demographic covariates. In terms of topics that are associated with running for office, topics 20 and 35 were also associated with rural and urban respondents, respectively, and refer to practical considerations progressive candidates would need to make in running for office in rural and urban districts, respectively. Topic 22, which was more prevalent among non-white respondents, deals with community and leadership. In terms of topics that are associated with *not* running for office, topics 33 and 36, which were associated with male respondents, correspond to frustration with the current political climate and Donald Trump, specifically.

Below, we show the words associated with some of the topics that were significantly associated with respondents who did not run for office. – topics 7, 9, 11, and 23. Some of these topics contain words suggesting vague, general interest in running for office (always, involved, time). A quick inspection of the documents most representative of topic 9, for instance, reveals relatively short statements of interest with sentiments such as "Would like to save the world from rightwing backlash" and "The world is on fire and I want to put it out while there's still something worth saving." Documents most representative of topic 11 indicate a general interest in politics and "making a difference" without mentioning specific offices or issues in which the respondent plans on making said difference. Finally, topic 23 is heavily weighted toward scientific terms, and includes a large number of soon-to-be or recent graduates of bachelors or masters programs (age is negatively associated with prevalence for topic 23, indicating that younger respondents were more likely to mention it). Previous research has shown that interest alone is seldom enough to convert ambition into an actual decision to run (Maestas et al. 2006). For respondents who mention these topics, this is the case.

```
##  TOPIC 7 TOP WORDS
##  ive, polit, now, alway, involv, time, get, never, ever, thought
##  TOPIC 9 TOP WORDS
##  like, feel, dont, think, lot, world, know, there, mayb, realli
##  TOPIC 11 TOP WORDS
##  make, can, impact, sure, know, place, communiti, decis, great, contribut
##  TOPIC 23 TOP WORDS
##  scienc, polit, degre, studi, master, current, scientist, scientif, engin
```

Next, we show the words associated with some of the other topics that are significantly associated with respondents who did run for office – topics 4, 5, 8, 18, and 29. These topics span a wide range of subjects. Topic 4 seems to be separated by the word "running" itself, with some respondents indicating that they are already running for office. Topic 5 deals with representation – either how the respondent plans to represent their constituents if they win or how the current representative for their district is not adequate. Topic 8 is marked by general sentiments regarding deservingness and opportunity, as well as a range of issues. Topic 18 is marked by notions of service, with respect to the public service of political office or past service such as time in the armed forces. Finally, topic 29 clearly relates to education, with respondents either speaking of the need for improved education/ education funding in their districts or their time in the classroom itself.

```
##  TOPIC 4 TOP WORDS
##  run, offic, think, year, interest, consid, decid, possibl, probabl, someday
##  TOPIC 5 TOP WORDS
##  repres, need, district, state, someon, candid, constitu, represent, someon, touch, desper
##  TOPIC 8 TOP WORDS
##  better, opportun, everyon, children, live, believ, deserv, ensur, access
##  TOPIC 18 TOP WORDS
##  communiti, serv, public, servic, work, advoc, organ, dedic, veteran, sector, servant
##  TOPIC 29 TOP WORDS
##  school, educ, student, public, board, teacher, colleg, high, classroom, english
```

As one candidate whose statement of interest was particularly representative of topic 8 wrote, "I am running to better North Dakota and create 21st century jobs. North Dakota has an oil boom that will dry up eventually and we are one of the leading states for natural resources such as Wind/Solar. I want to create 21st Century Jobs by harnessing those resources and building power grids to better our environment and create jobs for thousands of workers here in North Dakota."

And as a candidate whose statement of interest is representative of topic 18 wrote, "I served our country for over 8 years in the Marine Corps, and have served my home community for over 13 years. It's time to serve again." For these individuals, they have converted their political ambition into an actual decision to run. Although structural, personal, and political factors all contribute to mediate this conversion (Fulton et al. 2006), our data suggests that different motivations that contribute to political ambition matter as well for converting ambition into actually running for office.

PREDICTING CANDIDATE EMERGENCE USING STATEMENTS OF INTEREST

Lastly, we can use the proportion of words each document contains for each topic to re-predict candidate emergence using an elastic net, which is a form of regularized regression. As document-level topic prevalence was derived using demographic covariates, we do not include them in this model, instead using only information about topic prevalence in each respondents' statement of interest to predict whether they eventually ran for office.

Overall, the elastic net correctly classifies roughly 90 percent of the respondents. However, as less than 12 percent of the respondents on which the model is specified ran for office, this is not the best metric to use in order to evaluate performance. A demographics-only model, specified separately and alluded to above, correctly classifies roughly 88 percent of respondents by simply predicting that none of them run for office. Instead, here we look at the percent of each predicted classification (ran, did not run) that is correct, rather than the overall percentage of correct classifications.

This form of prediction accuracy is shown in Table 5 below. As it shows, of the 538 respondents who the model predicts ran for office based on what they said in their statement of interest, 396 (74 percent) did so. Of the 9845 candidates that the model predicted did not run for office based on what they said in their statement of interest, 813 (8 percent) did so. While this classification is by no means perfect, it is a dramatic improvement over a demographics-only model, indicating that the respondents' statements of interest provide meaningful information about their likelihood of converting from prospective to actual candidates.

TABLE 5 | **PREDICTIVE ACCURACY OF TOPIC PREVALENCE FOR CANDIDATE EMERGENCE**

	PREDICTED NON-CANDIDATE	PREDICTED CANDIDATE
Non-Candidate	9,053	142
Candidate	813	396
% Correct	**92%**	**74%**

That how prospective candidates articulated their interest in running provided more information about whether they would actually run than their demographic characteristics is worth emphasizing. Per our conversations with Run for Something, the organization provided support for any candidate who opted in to each successive stage of their recruitment process -- first attending a conference call, and then meeting 1-on-1 with a staffer or volunteer. That is to say, the organization did not systematically use the content on the intake forms in order to decide which prospective candidates to pursue. This being the case, we are reasonably confident that the reasons given for running were a *meaningful* and *independent* factor in separating those who eventually ran from those who did not.

CONCLUSION

This analysis is limited to prospective candidates who filled out a form with a specific candidate recruitment organization, so we would caution against using it to make broad generalizations about American political dynamics. However, we also note that a relatively large number of prospective candidates filled out Run for Something's intake form (roughly 10,000 forms had both an identifiable birth year and more than a handful of words in the statement of interest), making it one of the larger organizations of its kind in the country. This being the case, the data presented here can be taken as a useful case study in the current state of progressive organizing, mobilization, and candidate emergence. Our analysis has generated a handful of key findings:

- **ISSUES MATTER.** Respondents who eventually became candidates were more likely to mention specific issues they cared about.
- **REPRESENTATION MATTERS.** Women and people of color were particularly likely to discuss a need for better descriptive correspondence between the representatives and the represented in their communities.
- **AUGMENTING THE TRADITIONAL CANDIDATE PIPELINE.** Run for Something garnered interest from a variety of people who otherwise might not have approached their local party organization about running for office. While we found some similarities in candidate emergence patterns in the broader two-party system (age mattered more for men than women), we also found important differences (people of color accounted for a greater share of candidates than of respondents as a whole).
- **TALK ISN'T CHEAP.** What respondents said in their statements of interest was better at predicting candidate emergence than all of the demographic indicators available in the dataset.

Political observers have frequently attributed the unmistakable surge in interest in running for office over the course of the Trump presidency to Trump himself. However, while it may be the case that frustration with President Trump and his administration have driven prospective candidates to organizations like Run for Something, our analysis indicates that relatively few prospective candidates are articulating their interest in reference to the administration. Furthermore, those who do are significantly less likely to convert their interest into an actual candidacy. Instead, respondents who eventually became candidates were especially likely to discuss issues they care about (often health care and education), concerns regarding representation gaps, and matters that directly relate to their local communities. In short, the new wave of progressive candidates emerging in the wake of Donald Trump's election have something to run for.

REFERENCES

Alter, Charlotte. 2018. "A Year Ago, They Marched. Now a Record Number of Women Are Running for Office." http://time.com/5107499/record-number-of-women-are-running-for-office/.

Center for American Women and Politics. 2018. "2018 Summary of Women Candidates." cawp.rutgers.edu.

Fulton, Sarah, Cherie Maestas, L. Sandy Maisel, and Walter Stone. 2006. "The Sense of a Woman: Gender, Ambition, and the Decision to Run for Congress." *Political Research Quarterly* 52 (2): 235–48.

Maestas, Cherie, Sarah Fulton, Sandy Maisel, and Walter Stone. 2006. "When to Risk It? Institutions, Ambitions, and the Decision to Run for the U.S. House." *American Political Science Review* 100 (2): 195–208.

Schneider, Elena. 2018. "Millennials Take on Trump in the Midterms." https://www.politico.com/story/2018/05/28/democrats-millennials-midterm-elections-610024.

Thomas, Sue. 2002. "The Personal Is the Political: Antecedents of Gendered Choices of Elected Representatives." *Sex Roles* 47: 343–53.

Tumulty, Karen. 2018. "There's an Election Wave You Probably Haven't Heard of yet – and It's Important." https://www.washingtonpost.com/opinions/theres-an-election-wave-you-probably-havent-heard-of-yet--and-its-important/2018/05/30/7854fd4e-6344-11e8-99d2-0d678ec08c2f_story.html?utm_term=.370c9c969eb1.

Yong, Ed. 2018. "Here's How the Scientists Running for Office Are Doing." https://www.theatlantic.com/science/archive/2018/02/heres-how-the-scientists-who-are-running-for-office-are-doing/554444/.

ENDNOTES

1. Run for Something omitted personally identifying metadata, such as names, from the data provided to Data for Progress. In cases where the statements of interest contained information that could be used to identify respondents, Data for Progress has omitted it from published results.

2. Top words were determined based on their probability of appearing in each topic and their FREX score, which is a metric that identifies words that are both frequent and exclusive to the topic.

ABOUT THE AUTHORS

MEREDITH CONROY is an associate professor of political science at California State University, San Bernardino. Her work has appeared in *Politics, Groups, and Identities* and *American Politics Research*. Her first book, *Masculinity, Media, and the American Presidency*, explores the extent to which the media plays a role in the maintenance of masculine gender norms in U.S. presidential politics.

JON GREEN is a Ph.D. student in political science at Ohio State University. His work has appeared in *Science and Perspectives on Politics*.

ACKNOWLEDGEMENTS

The authors thank Run for Something for sharing their data, and specifically thank Sarah Horvitz for helping organize and understand it. The authors also thank Billie Kanfer for designing this document; Erin Cassese, Jacob Smith, and Jon Kingzette for helpful feedback; and Sean McElwee for institutional support.