How do female candidates affect voter turnout?

Daniel B. Jones[a]
University of South Carolina

Abstract. How does the presence of a woman on the ballot impact election outcomes, aggregate turnout, and the voting behavior of particular groups of voters? Using a regression discontinuity approach, I exploit quasi-random variation in the presence of a female candidate in US House elections stemming from narrowly won primary elections between candidates of different genders. I find that the presence of a female candidate leads to lower overall turnout, but otherwise has no bearing on the outcome of the election. The change in turnout is driven entirely by male voters, which falls uniformly for (male) voters of both parties.

[a] University of South Carolina, Darla Moore School of Business, Department of Economics, 1014 Greene St., Columbia, SC 29208. Email: daniel.jones@moore.sc.edu.

1. Introduction

Does the gender of a candidate impact turnout and voter behavior? There are many reasons - anecdotal and theoretical - to think there might be a relationship between candidate gender and voter behavior. Anecdotally, whenever a female candidate is in a race for a major office, we see countless articles in the popular press speculating on how women will vote as a response (e.g., "Is Fiorina GOP's magic key to the women's vote?"[1], "No Banking on the Women's Vote**:** Hillary Clinton will have to be mindful that each generation must be inspired by the promise of a better tomorrow"[2]). Theoretically, there is some ambiguity as to how candidate gender may impact voter behavior. An implication of the Downsian model of electoral competition, wherein candidates converge to the preferences of the median voter, is that identities of candidates (e.g., gender) have no bearing on voter behavior. Other models (for instance, the citizen-candidate model) assume candidates are motivated to implement their preferred policies (Osborne & Slivinski, 1996; Besley & Coate, 1997). If women have different policy preferences and adopt those preferences as their platform[3], then – at a minimum – we would expect to see a change in turnout patterns. Women, in particular, may be more likely to turn out as the policy being proposed may also be closer to their own preferred policy. Finally, in the presence of gender bias, turnout may decline, particularly among male voters (existing work on this possibility is reviewed below.)

With these issues in mind, in this paper I ask: How does the presence of a woman in an election impact the election's outcome? How is turnout impacted? Are particular groups of

[1] http://www.seattletimes.com/nation-world/is-fiorina-gops-magic-key-to-womens-vote/
[2] http://www.usnews.com/opinion/blogs/penny-lee/2013/04/04/hillary-clinton-shouldnt-count-on-the-womens-vote-in-2016
[3] Several researchers have documented that women's suffrage in the US measurably altered policy outcomes (Lott, 1999; Carruthers and Wanamaker, 2015). A number of papers have found that the election of a woman has an impact on policy outcomes (Anzia & Berry, 2011; Clots-Figueras, 2012; Bhalotra & Clots-Figueras, 2014).

voters, especially women, more likely to turnout? Are male voters less likely to turnout? To address these issues empirically, I use a regression discontinuity strategy that quasi-randomly assigns a woman (rather than a man) to run as a candidate for a seat in the US House of Representatives. I use data on election returns from the primary elections that established a candidate's presence in the general election. In primaries where the two leading candidates are of opposite genders, a sufficiently narrow primary race quasi-randomly assigns whether a female candidate is involved in the general election. Using vote shares from the general election, I examine whether a party's candidate fares better or worse when that candidate is a woman. Ultimately I find that the presence of a female candidate leads to lower turnout. This, however, does not impact the outcome of the election as measured by candidates' vote shares; I find no evidence that a female candidate receives a clear advantage or suffers any clear disadvantage in the general election.

To better understand the mechanisms driving the main results, I draw on voter-level survey data (the *American National Election Survey*); these data allow me to consider how different types of voters react to the presence of a female candidate. Analyses of the survey data yield several findings. First, female voters are generally unresponsive to the presence of a female candidate; they are neither more nor less likely to turn out when a woman is on the ballot. Men, on the other hand, are significantly less likely to turn out. This is true whether they identify with the same political party as the female candidate or the opposing party. This, therefore, explains how turnout can fall without impacting the outcome of the election: men of both parties are uniformly less likely to vote when a female candidate is involved.

This leaves the question of *why* turnout declines when female candidates are involved in the race. Additional analyses suggest that both candidates (the randomly-assigned candidate and

his or her opponent) campaign with a more ideologically extreme position when a female candidate is involved in the race; that is, when the Democratic candidate is female (rather than male), the Democratic candidate adopts a more extreme left position and the opponent (a Republican) adopts a more extreme right position than would be the case in a race between two male candidates. While I cannot claim to perfectly identify the mechanism driving my results, these patterns suggest alienation of centrist voters as one possible explanation for the decline in turnout from both parties' voters (consistent with theoretical models of electoral competition tha allow for abstention (Anderson & Glomm, 1992; Llavador, 2006)).

The existing literature has studied the impact of female candidacy on turnout in one of two ways: (1) exploiting variation in presence of female candidates resulting from the introduction of gender quotas on party lists, or (2) surveys and/or laboratory experiments. De Paola et al. (2014) and Casas-Arce and Saiz (2015) are recent examples of the first type of study. Both find some evidence of a causal positive impact of female candidacy (driven by gender quotas) on turnout. Relative to those studies, this paper contributes to the literature by considering an exogenous shock to female candidacy that is not driven by a policy change. When the frequency of female candidates increases due to gender quotas, it is at least possible that voters are reacting to the policy rather than to the candidates themselves. Indeed, my findings suggest that turnout (especially male turnout) declines when a female candidate is present in an otherwise "ordinary" election.

Wolak (2014) conducts a laboratory experiment wherein she presents participants with hypothetical candidate information, but experimentally varies whether participants believe one, both, or neither of the candidates are female. She finds that the presence of women in an election does nothing to increase women's reported likelihood of voting in the hypothetical scenario and

decreases men's likelihood of voting. This is not unprecedented; research using survey data also finds that greater female representation has little impact on women's political engagement (Lawless, 2004; Dolan, 2006).[4] In short, the hypothetical and/or survey-based link between female candidacy and female voter mobilization is tenuous. I contribute to this literature by providing a natural experiment to assess the causal impact of female candidacy on actual voting behavior, and ultimately find results that are consistent with these existing findings, especially those of Wolak (2014).

Although the main focus of this paper is the impact of female candidacy on turnout, the question of whether female candidacy impacts election outcomes (e.g., the relative vote shares of candidates) is explored as well. Existing research using experiments, surveys, and election returns finds no clear evidence that the gender of a candidate (holding other candidate and policy characteristics constant) influences voters (e.g., McElroy and Marsh, 2009; Black and Erickson, 2003). In a very recent study, Anastasopoulos (2016) draws on election returns data to assess the impact of candidate gender on vote share and likelihood of winning; he finds no evidence of a "gender penalty." I ultimately find similar results in this paper. I build on Anastasopoulos's findings, exploring the impacts of a female candidate on turnout, both in aggregate and decomposed by subgroups of voters (e.g., female voters, voters of the same party as the female candidate).

Anzia and Berry (2011) suggest that female candidates may require a higher level of qualification (e.g., prior legislative experience) in order to counteract gender bias and attain the same vote share that a less-qualified male candidate would.[5] This argument motivates some of

[4] More recent work highlights the importance of considering the interaction of gender and party congruence between voter and female candidate, which motivates some of the analysis in this paper (Reingold and Harrell, 2010).
[5] To support the claim, they document that women elected to Congress are more successful in attracting federal funding to their districts and sponsoring legislation than men.

the validity checks and controls I include in analyses in later sections. Related to this, Anastasopoulos (2016) argues that, if female candidates were required to be "higher quality" candidates to compete in the context of US House elections, we might expect that their fundraising abilities would exceed those of male candidates. Using data on campaign donations, he finds no evidence that this is the case.

2. Data and empirical approach

To assess the causal impact of a female candidate on election outcomes, I use a regression discontinuity approach taking advantage of close elections between male and female candidates at the primary election stage. These close elections generate quasi-random assignment of candidates to the general election. I focus on elections to the US House of Representatives between 1972 and 2010. To fix the terminology that I will use throughout, I will refer to the candidate who narrowly wins a primary between a man and a woman as the "randomly assigned candidate". I will refer to the candidate he or she faces off against (who is *not* randomly assigned to participate in the general election) as the "opposing candidate". The main outcome variables of interest include the "randomly selected" candidate's vote share and voters' likelihood of turning out in "treated" general election races.

2.1 Data

The data I use on US House primary election returns were assembled by Pettigrew et al. (2014). These data include the universe of primary election candidates (from every congressional district, party, and election cycle) during my sample period. Critically, these data include the primary election vote share and gender of each candidate. Because I use these data to randomly

assign whether a male or female participates in the general election, I ultimately only maintain the subset of primary elections with at least two candidates where the top two primary candidates (the winner and the first loser) are of opposite genders. Of 6,248 primary elections with at least two candidates during the sample period, roughly 20% (1,246 unique elections) feature one man and one woman in the top two positions. (Only 1.4% of these elections feature two women competing to win the primary.) Of the primary elections with one man and one woman, 61% are Democratic primaries. I further restrict my sample to general elections where the opposing candidate is a man. This is true of 88% of observations in my sample.

Of course, in the regression discontinuity design, identification is drawn from the narrow elections. Thus, it is worth noting that there is a relatively large number of narrowly determined primaries in my sample. Of the 957 elections in my estimating sample, 208 of the treatment-inducing primary elections are decided by a margin of 10 percentage points or less; 106 of the elections are decided by a margin of 5 percentage points or less; 56 of the elections are decided by a margin of 2.5 percentage points or less.

I combine the primary election returns data with two main sources of data used here as outcomes: (1) US House general election returns from 1972-2010, drawn from the CQ Press Voting and Elections Collection and (2) voter-level survey data (the *American National Election Studies* (ANES)). Additional data on outcomes are drawn from the Database on Ideology, Money in Politics, and Elections (DIME), which are available only from 1980-2010 (Bonica, 2013).

The CQ Press election returns report the share of the vote received by each candidate participating in general elections during the sample period for all US House elections. The data also report the total number of votes cast in each election. I use the vote share of the randomly assigned candidate as one of the main outcome variables in the paper; this allows me to assess

whether a female candidate face an electoral advantage or disadvantage, and whether the presence of a female candidate impacts the outcome of an election more generally. I also assess how the presence of a female candidate impacts turnout. The CQ Press data report the total number of votes cast, but not the number of people in the district registered to vote. Thus, I pair this with Census data measuring the total population of the Congressional district in each decade. I therefore present two measures of turnout: (1) the number of votes cast in the district divided by the population of the district (votes per capita) and (2) the raw vote total within the Congressional district, with a control for the total population in the district in the regression analysis.

The ANES data is based on a survey. ANES surveys a random sample of voters each federal election year. Importantly, the ANES identifies the congressional district that a voter lives in, so I can match respondents in the ANES data to the primary data to determine whether a respondent is "treated" (that is, experiences a narrow victory of a female in one of the district's primary). Many questions asked in the survey change from year-to-year (e.g., questions on voter attitudes towards specific issues), but the limited set of questions I use here remain fairly constant. Specifically, in the ANES, respondents are asked whether they voted in the most recent November general election. This is the second main outcome variable I use in the paper. Respondents also report political leanings (Republican lean, independent, Democratic lean) and provide limited demographic information including gender, race, and age. Using these data (combined with the primary election 'treatments'), I can assess whether respondents are more or less likely to vote when a female candidate is involved. I can also assess how any changes in turnout vary by gender and/or party match or mismatch with the randomly-assigned candidate.

I also use one additional source of data (the *Database on Ideology, Money in Politics, and Elections*, or "DIME") to provide some supplementary results on candidates' ideological positioning during the campaign. Bonica (2013) compiled campaign finance information for all Congressional candidates from 1980-2010 to construct a novel ideological positioning measure for candidates based on the ideological positions of their donors. This measure is called the "CF-Score," and like other such measures, places candidates on a left-right continuum where increasingly liberal (conservative) candidates are assigned increasingly negative (positive) scores. A major advantage of this measure relative to more common ideological measures (e.g., DW-Nominate scores) is that every *candidate* has a score, even if they never serve in office. Thus, I can assess whether the presence of a woman in a Congressional election leads candidates to adopt different positions than they would otherwise.[6] There are two versions of the CF-Score; one is constant throughout a candidate's career, the other (the "dynamic CF-score") changes with each campaign cycle. I use the dynamic score, as it allows me to assess how candidates position changes given the opponent they face in a given campaign.

Finally, as Anzia & Berry (2011) note, female candidates may require higher qualification for office in order to attain the same vote share as a less-qualified man. In an attempt to address this issue, I employ controls for prior legislative experience of the randomly assigned candidates in some specifications. These controls are drawn from two sources. An indicator for whether the randomly assigned candidate is an incumbent (and therefore has prior federal legislative experience) is included Pettigrew et al.'s (2014) primary election data. I also

[6] Notably, Bonica (2014) documents that, when both scores are available, the correlation between CF-Scores and the more commonly used DW-Nominate measure is very high. For full details about the CF-Score, see Bonica (2014).

obtain an indicator of prior experience in state legislatures. This is done by drawing on Klarner et al.'s (2013) state legislative election returns data.[7]

2.2 Empirical approach

Throughout the paper, I employ a local linear regression discontinuity approach with optimal bandwidths selected in accordance with Calonico et al. (2014). I focus on elections where one of the candidates emerged from a primary where the top two candidates were of opposite genders. In sufficiently close primary elections, the candidate (male or female) is plausibly randomly assigned to participate in the general election.

I estimate variations of the following equation:

$$\begin{aligned} y_{dt} \quad = \quad & \propto + \beta_1 \mathbf{1}[Female\ wins_{dt}] + \beta_2 Female\ margin_{dt} \\ & + \beta_3 \mathbf{1}[Female\ wins_{dt}] \times Female\ margin_{dt} \\ & + \delta_t + \varepsilon_{dt} \end{aligned}$$

where the subscript "d" indicates the congressional district and subscript "t" indicates the election year. The variable "Female wins" is an indicator that is equal to one in districts and years where a female candidate won in a primary against a male candidate; that is, the randomly assigned candidate is a woman. "Female margin" represents the difference between female primary candidate's vote share and male primary candidate's vote share. "Female wins", "Female margin", and the interaction of the two mirror the relatively standard regression discontinuity approach. Analyses are restricted to samples wherein the female margin of victory

[7] ICPSR Study No. 34297. Those data include the names and electoral outcomes of all candidates for all state legislative elections in the United States from between 1967-2010. I use candidates' names and states to match those data to the randomly assigned candidates in my estimation sample in order to determine whether a candidate has previously been elected to their state legislature.

is within the optimal bandwidth, which varies depending on the dataset being used. (Additional analyses document robustness to varying the bandwidth.) δ_t are year fixed effects, which are important in this context to account for whether a given election is occurring during a presidential election.[8] In all models, standard errors are adjusted to allow for clustering at the state-by-decade level (where decades are defined as starting in years ending in "2" (e.g., 1992) to reflect the timing of redistricting.)

Because "Female margin" is the difference between the vote share received by the female candidate and the vote share received by the male candidate in the primary, β_1 can be interpreted as the impact of a female primary victory (and therefore general election candidacy) when the margin was zero; in practice, this is interpreted as the impact of a victory in a very close primary election. Thus, the "Female wins" coefficient is of primary interest in this specification, as it can be interpreted as the causal impact of a female candidacy.

The outcome variable y_{dt} captures the outcomes discussed in the previous section. In the election level analysis, which draws on the CQ Elections data, there are two main outcomes. The first is the randomly selected candidate's vote share in the analysis of general election returns. Vote share is defined as (Randomly selected candidate's votes)/(Total votes cast) and therefore runs from 0 to 1. (In almost all elections, either there are just two candidates or third-party candidates receive a very small number of votes; thus, a higher vote share received by the randomly selected candidate directly implies a lower vote share received by the opposing candidate.) The second main outcome is turnout. While I have a count of the total population in each Congressional district for every election in my sample, I do not have a measure of total *registered voters* by Congressional district. Thus, I measure turnout in two ways. The first is

[8] Results are similar excluding year fixed effects.

"number of votes cast per capita", or (Total votes cast)/(Total population in the district). In other specifications, I take (log of) total votes cast in the district as the outcome variable and simply control for (log of) total population.

To account for the concern raised by Anzia & Berry (2011), in some specifications, I add controls for randomly assigned candidates' prior legislative experience (as discussed in the data section). In those specifications, I also control for incumbency status of the opposing candidate.

Finally, in the voter level analysis (drawing on the ANES survey data), the main outcome is a dummy variable indicating that a respondent voted in the general election. This takes on the value 1 if the respondent voted and zero otherwise. The ANES data report information on voter characteristics (race, age, gender, and party affiliation), which I include as controls.[9]

As noted, the optimal bandwidth (selected using the Calonico et al. (2014) procedure) varies depending which dataset I am drawing on. In the election-level analysis (CQ data), the sample is restricted to elections where the treatment-inducing primary was decided by a margin of less than 23 percentage points; in the voter-level analysis, the sample is restricted to observations where the primary was decided by a margin of less than 19 percentage points.

Table 1 draws on the election returns data to provide a summary of characteristics of the elections in my sample. Column 1 provides summary statistics for the full sample of elections where in one party's primary involved candidates of opposite genders, and therefore at least had the possibility to randomly assign either a man or woman to participate in the general election. Columns 2 and 3 restrict the sample to the bandwidth selected through the Calonico et al. (2014) procedure and half of that bandwidth, respectively. Column 2 therefore summarizes the set of elections that form the sample in most of the regressions to follow. It is worth noting that all geographic regions and both major parties are relatively well represented in the sample. Also

[9] Results are similar without controls.

worth noting is that – compared to the full sample in Column 1 – the "randomly selected" candidates (that is, candidates that emerge from a primary between a man and a woman) in Columns 2 and 3 are weaker candidates: they are less likely to be incumbents, they are less likely to have prior state legislative experience, and they receive lower vote shares in the general election. This is not surprising: the narrow bandwidth is selecting candidates who narrowly emerged from their party's primary, whereas Column 1 includes many candidates who easily won their primary (winning by margins of greater than 23%) and are therefore revealed to have strong support within the district. Thus, it should be noted that the results that follow are based on a subsample of candidates who are in a weaker than average position in their general election race.

3. Empirical results

3.1 Simple graphical analysis

I begin by previewing two of the main results of the paper in simple graphical form, using only raw data. Figure 1(a) is a binned scatterplot plotting the relationship between the female candidate's primary election margin of victory (x-axis) and the general election vote share received by the candidate that ultimately emerged from that primary, or the "randomly selected candidate" (y-axis). A positive female margin of victory implies that the female candidate won the primary and is involved in the general election. When margin of victory is close to zero, the primary election was close. There is no evidence of a discontinuity at the cutoff; that is, it appears that there is no clear advantage that randomly-assigned female candidates receive (as might be expected if female candidates mobilized female voters with no parallel response from the opposing voters), nor do they experience a clear disadvantage (which

might result, for instance, from bias against female politicians). Of course, it is possible that some voters are more likely to vote for the female candidate, while others are less likely to vote for that candidate, and the two effects cancel each other out. This cannot be explored in these data, but will be addressed in a later section using the ANES survey data.

Figures 1(b) and 1(c) take two different outcomes variables capturing turnout on the y-axis: votes per capita and raw total votes. This allows us to assess whether the presence of a female candidate mobilizes (or demobilizes) voters, even if there is no impact on the outcome of the election. In both figures, there is evidence that the random assignment of a female candidate to the general election reduces turnout; both measures of turnout experience a discontinuous negative jump at the cutoff. These results are further probed with regression analysis in the remainder of the paper.

In the appendix, I report the results of standard regression discontinuity validity tests. To be brief: I find that the density of the running variable (female primary margin of victory) is smooth around the cutoff (Appendix Figure A1). If this were not true, we would be concerned that some unobserved factor provides female candidates an advantage or disadvantage in narrow elections that could compromise the validity of the design. Related to that concern, I plot a series of figures similar to Figures 1(a), (b), and (c) (and report results from regressions) to ensure that other relevant observables are smooth around the cutoff (Appendix Figure A2). The figures reveal no evidence that the narrow victory of a female is correlated with some other characteristic (including, for instance, likelihood of prior state legislative experience).

3.2 Regression results: The impact of a female candidate on election outcomes

I now turn to the main regression results of the paper. In this subsection (as in Figures 1(a), (b), and (c)), the main outcome variables of interest are the vote share received by the randomly selected candidate and overall turnout.

In Table 2 (Panel A), I report the results of the local linear regression discontinuity approach discussed in subsection 3.2. I report only the "Female wins" dummy, which provides evidence on the causal effect of a female candidate's presence in the race. Columns 1 and 2 take the randomly selected candidate's vote share as the outcome variable. Column 1 includes only year fixed effects (in addition to the "female winner" dummy, "female primary margin of victory" measure, and the interaction of those two variables). Column 2 adds controls meant to account for the candidate's qualifications (incumbency status and state legislative experience, as well as the opposing candidate's incumbency status).

In both columns 1 and 2, the estimates provide no evidence that female involvement in a Congressional election has a causal impact on the outcome of the election; the estimated coefficient is very close to zero. That is, there is no clear evidence that female candidates receive a clear electoral advantage (which would yield a positive coefficient in columns 1 and 2) or a clear disadvantage (which would yield negative coefficients). It is of course possible that there is some combination of positive effects (e.g., female mobilization) and negative effects (e.g., gender bias) occurring simultaneously, but these data do not directly speak to the mechanism driving the result. Later in the paper, I analyze survey data from individual voters to further explore what drives the absence of an effect of female candidacy on vote shares.

While female candidacy does not impact the election outcome, it does appear to impact turnout. This can be seen in Columns 3-6. Columns 3 and 4 take "votes per capita" as the outcome variable. (Recall that this is simply "total votes" divided by "total population within the

district," as I do not have registered voter counts.) Columns 5 and 6 are more flexible; they take the log of total votes as the outcome and control for log of population on the right hand side. Columns 4 and 6 include the same additional controls that are included in Column 2.

In short, all four columns reveal that turnout falls when a woman is involved in a Congressional election (albeit at only a 10% level of significance in three of the four columns). Columns 3 and 4 suggest that the fraction of the total population that votes falls by roughly 2 percentage points (keeping in mind that a large fraction of the population is either not eligible or not registered to vote). Columns 5 and 6 suggest that the total number of votes cast is roughly 8% lower in elections where a female is involved as compared to contests between two men.

Two tables in the appendix assess the robustness of these findings to alternative specifications. Appendix Table A2 reports results taking either half of the optimal bandwidth or double the optimal bandwidth in the local linear regression approach. Appendix Table A3 adopts a global parametric regression discontinuity approach (as in Ferreira & Gyourko (2009, 2013)). Both sets of alternative specifications yield similar results.

Table 2 (Panel B) tests for heterogeneity in the findings thus far. Specifically, I assess whether the results differ when the randomly selected candidate is a Democrat or a Republican. For instance, one party may favor female candidates, while the other opposes them; this would make it seem as though vote share is unchanged with a female candidate, where in reality there are effects in opposite directions depending on party. To test this, I modify the main estimating model to allow for an interaction between "female winner" and the party of the candidate that narrowly emerged from a primary. The coefficient on "Fem Wins X Dem" provides evidence on the causal impact of a Democratic female being involved in an election; "Fem Wins X Rep" provides evidence on the causal impact of a Republican female being involved in an election.

Columns 1 and 2 of Table 2 (Panel B) reveal no evidence of heterogeneity by party in the vote share result; regardless of the randomly selected candidate's party, there is no clear electoral advantage or disadvantage for female candidates. There is, however, heterogeneity by party in the impact of a female candidate on turnout (Columns 3-6); turnout falls specifically when a Democratic female candidate is involved in the race. The presence of a Republican female has no clear impact on election outcomes or turnout.[10]

3.3 Empirical results: The impact of a female candidate on individual voter turnout behavior

Having documented that the presence of a female candidate has no impact on aggregate electoral outcomes but does seem to reduce turnout, I now explore individual-level voter turnout to better understand why that might be the case. The results in this section use the ANES survey data drawn from the same elections and years as I studied in the previous subsection, with the exception of 2010 for which there is no ANES data. Table 3 provides summary statistics for the ANES variables and sample used in this section. The main outcome variable is whether an individual reports voting in the November general election for their US House representative, which equals 1 if so and 0 otherwise. It is worth noting that the self-reported rate of voting is rather high. Thus, I do not intend to interpret the magnitudes of the coefficients in the regression analysis.

In Column 1 of Table 4, I employ the same local linear regression discontinuity approached used in the previous section, now taking the indicator variable capturing whether the respondent voted in their House election as the outcome. Recall that the ANES data indicate

[10] It is interesting to note that this, to some extent, parallels Washington's (2006) findings on the impacts of black candidates. She finds that only Democratic black candidates impact turnout. Voters are unresponsive to the race of Republican candidates, just as they are seemingly unresponsive to the gender of Republican candidates in the findings presented here. (However, unlike the findings I present on female candidates, she finds that black candidates *increase* turnout.)

which Congressional district a respondent lives in, so we can observe whether they are "treated" and whether they vote in the relevant general election. In these data, the bandwidth around the cutoff is 19%. As before, I only report the coefficient on the "Female wins" dummy, which provides evidence on the casual effect of a female candidate's presence in a general election. Other controls are included as noted in the data section.

As in the analysis of aggregate turnout data, Column 1 of Table 4 suggests that the presence of a female candidate leads to a lower likelihood that an individual voter turns out. This result, however, is not statistically significant at conventional levels. In Appendix Table A4, I re-run this specification without controls and with different bandwidths. The same general picture emerges there: a female candidate has a consistently negative (but imprecisely estimated) impact on likelihood of voting.

Recall, however, that the reduction in aggregate turnout was primarily driven by elections where the randomly selected candidate was a Democrat. In Column 2, I test whether the same is true in these data (which could explain the imprecise negative in Column 1). To do so, I adopt the same approach used in the previous section; specifically, I interact the treatment indicator ("Female wins") with dummies indicating whether the treatment-inducing candidate is a Democrat or Republican and fully interact the resulting indicators with the margin of victory. Indeed, while there is small and insignificantly positive effect of a female candidate on turnout when that candidate is a Republican, a Democratic female candidate leads to significantly lower reported turnout amongst survey respondents.

In Column 3, I test for heterogeneous responses to female candidates by gender; this was not possible in the aggregate turnout data and represents an advantage of the ANES data. I fully interact "Female wins" with the gender of the respondent to identify the causal impact of a

female candidate on male and female voters. Ultimately, the decline in turnout observed throughout the rest of the paper appears to be driven by men. There is no detectable impact of a female candidate on a female respondent's likelihood of voting in their house election.

Next, I further probe these results to understand the decline in turnout that is seemingly driven by male voters and situations where the female candidate is a Democrat. Specifically, I aim to better understand whether the preceding results depend on whether the female candidate is affiliated with the party that the voter prefers ("party congruence") or not. This will shed light on the mechanism driving the results. For instance, if gender bias were the primary driver of the decline in turnout (especially amongst men), one might expect this result to only occur if the voter would typically vote for party of the female candidate (but is now less likely to as a result of the candidate's gender). One might even expect turnout amongst gender biased voters to be *higher* when the candidate from the *opposing* party (relative to the voter's own preferences) is a female, as they turn out to "vote against" the female candidate. Alternatively, a preference for female candidates (whether specifically amongst females or not) may lead to the opposite result: female-preferring voters may be more likely to turnout when the candidate is from their own party, but less likely to turnout when the female candidate is from the opposing party (because they would rather not vote *against* a female candidate).

With these issues in mind, Table 5 assesses the impact of party congruence with the female on the likelihood of turning out. Again, I fully interact the "Female wins" dummy (and the "female primary margin of victory") with dummies indicating that either: (1) a voter's partisan leaning is the same as the party of the randomly assigned candidate ("Voter Party = Cand. Party") or (2) a voter's leaning is *not* the same as the party of the randomly assigned candidate ("Voter Party ≠ Cand. Party"). Column 1 reports the results using the same estimating

sample that is used in Table 4; there is no immediate evidence that the impact of a female candidate depends on party congruence in the full estimating sample. The remaining columns of Table 5 split the sample either by party of the randomly selected candidate (Columns 2 and 3) or by the gender of the survey respondent (Columns 4 and 5). Column 2 again reveals that a female candidate on the Republican side has no (statistically significant) impact on turnout (as in Table 4), but further shows that this for both Republican-leaning voters and Democrat-leaning voters. Column 3, on the other hand, shows that the decline in turnout that occurs when a (randomly-assigned) female candidate is a Democrat is driven by a reduced likelihood of both Democratic and Republican voters turning out. This helps reconcile two results documented thus far: when a (Democratic) candidate is female (1) turnout is lower and (2) the election outcome (as measured by relative vote share for the randomly assigned candidate) is unaffected. We now see that this occurs because voters from *both* parties are less likely to turnout when a female candidate is in the race; the reduced turnout from both sides prevents the female candidate from experiencing any clear advantage or disadvantage.

Columns 4 and 5 assess the interaction of voter gender and voter-candidate party congruence. In short, the conclusions drawn from Table 4 (that female voters are unaffected by female candidates while male voters are less likely to turn out) are also unaffected by allowing for heterogeneity in party congruence. Men from both parties (the party of the randomly-assigned candidate *and* the opposing party) are less likely to turn out when the randomly assigned candidate is female.

To summarize: the presence of a female candidate reduces voters' likelihood of turning out. This is true in particular when the candidate is a Democrat, or when the voters are male.

These results do not depend on the alignment of partisan leaning between voter and candidate; *all* voters are less likely to turn out when the candidate is a woman.

What does this combination of results imply? While I cannot claim to clearly identify the mechanism driving changes in turnout, the results are at least suggest that the decline in men's turnout is not driven by gender bias—particularly given that men from both parties are less likely to turn out when one candidate is a woman. I draw on a different outcome variable in the ANES data to provide additional (albeit merely suggestive) evidence on this point. Specifically, respondents are asked whether they "like" candidates running for Congress in their district. In Appendix Table A5, I assess whether respondents are more or less likely to report positive feelings about the randomly-assigned candidate when that candidate is a woman; as before, I test for heterogeneity by party of candidate, party congruence with candidate, and respondent gender. The outcome variable is equal to one if respondents report positive feelings about the randomly-assigned candidate and zero otherwise. The same local linear regression discontinuity approach used throughout the paper is employed here as well. (Note that this question was not asked in all of the years used in the rest of the analysis, so the sample size is smaller.)

Column 1 reveals some positive relationship between a female candidacy and respondents' subject impression of the candidate, but this relationship is not statistically significant. Column 2 shows that this is mostly driven by female candidates who are Democrats. This is, of course, somewhat surprising given that it was only amongst Democratic candidates that a female candidacy led to *lower* turnout. Column 3 reveals that candidates are slightly more likely to report positive feelings about a positive candidate with whom they agree. Column 4 documents very little difference between male and female respondents' feelings about a female candidate; both are (insignificantly) more likely to report positive feelings when the candidate is

female. This is noteworthy in light of the gender bias possibility. While it is men (from either party) who are less likely to vote when one candidate is a woman, they are not less likely to report positive feelings about that candidate (and in fact may be slightly more likely to do so). Obviously, this should be interpreted with caution, as respondents have no incentive to report their true opinion of the candidate. Nonetheless, the data we have on this issue -- paired with the fact that men from both parties are less likely to turn out -- do not point towards clear evidence that gender bias drives the decline in turnout.

3.4 Empirical results: The impact of a female candidate on candidates' ideological positioning

The results thus far suggest that the presence of a female candidate leads to lower turnout, especially when that candidate is a Democrat. This is driven by a decline in turnout of male voters from both parties. Because turnout from both parties declines, the outcome of the election (as measured by relative vote shares) is unaffected. What explains the decline in turnout? While I cannot rule out gender bias altogether, results from the previous subsection point away from that being the main explanation.

In this subsection, I explore a different possibility: the nature of the campaign may change when a female candidate is involved. Specifically, candidates may be pushed to different ideological positions than they otherwise would be, which may alter voters' interest in turning out. To assess this, I draw on Bonica's (2013, 2014) dynamic "CF-Scores", a measure of ideological positioning of candidates during a particular campaign cycle. This measure places candidates on a left-right spectrum based on the the ideology of their donors during the

campaign. Increasingly negative (positive) numbers indicate that the candidate's ideological position is further to the left (right) during a given campaign.

Table 6 repeats the same local linear regression discontinuity approach that has been used throughout the paper, but takes either individual candidates' CF-Scores or the divergence between candidates as outcome variables. Because the CF-Score data is available for a smaller number of years (1980-2010), the optimally selected bandwidth is slightly larger than it was in previous subsections of analyses; here, the bandwidth is roughly 0.31. I report the means of the outcome variables used in this subsection in the top panel of the table.

Panel A reports results for elections where the randomly assigned candidate (that is, the candidate who narrowly emerged from a mixed-gender primary) is a Republican. Recall from previous subsections that there was little or no impact of a Republican candidate's gender on turnout; the decline in turnout was almost entirely driven by female Democratic candidates. Here, we see there is also no impact of a female candidate (who is Republican) on ideological positioning during the campaign. Neither the Republican candidate (Column 1) nor the opposing Democrat candidate (Column 2) have different ideological positions when the Republican is female than they would hold if both candidates were male. As a result, the divergence between the candidates (measured as the difference between the Republican's score and Democrat's score in the relevant election) is not affected by the presence of a female candidate (Column 3).

Panel B reports results for elections where the randomly assigned candidate is a Democrat. There, we see a clear impact of female candidacy. When the Democratic candidate is a female, both the Democratic candidate and the Republican candidate adopt more extreme positions during the campaign than they would if the Democratic candidate were male. The randomly assigned Democrat is pushed further to the left (Column 2), and the opposing

Republican candidate is pushed further to the right (Column 1). This leads to a relatively large increase in the divergence between candidates (Column 3), increasing the gap between candidates' positions by nearly two-thirds of a standard deviation.

Theoretical models of electoral competition that allow for voter abstention highlight that, when voting is costly, increasingly extreme candidates may alienate moderate voters and therefore lead to lower turnout (Anderson & Glomm, 1992; Llavador, 2006). Drawing from these models and the pattern of empirical results documented here, the increased policy divergence between candidates when a female candidate is involved in the race may explain the decline in turnout.

4. Conclusion

In this paper I assess the relationship between the presence of a female candidate, turnout, and election outcomes. To empirically test the possible impacts of a female candidacy, I exploit quasi-random assignment of female candidates in general elections for the US House stemming from narrowly won party primary elections between a man and a woman. Ultimately, I find that the presence of a female candidate has no impact on the outcome of the election; this is consistent with work that has tested for a "gender penalty".[11] This does not imply that voters are oblivious to gender. The presence of a female candidate reduces turnout, both in aggregate elections returns data and individual-level survey data.

In the voter-level data, I explore how particular groups of voters (by gender and partisan affiliation) respond to the presence of a female candidate to probe what drives the patterns in the aggregate data. The data suggest that men in particular are less likely to vote when one candidate is a female; female voters are non-responsive to gender of candidates. The decline in turnout that

[11] McElroy and Marsh, 2009; Black and Erickson, 2003; Anastasopoulos, 2016

occurs when a female candidate is involved in the race occurs relatively uniformly across voters of both parties, thereby explaining how turnout could fall without impacting candidates' relative vote shares. This fact (paired with data on voters' subjective opinions of candidates) point away from gender bias as an explanation for the reduced likelihood of turning out.[12] Instead, drawing on measures of candidates' ideological positioning during the campaign, I find that the policy divergence between candidates is larger when a female is involved in the race (driven by more extreme positioning of both the female candidate and her opponent). This, rather than gender bias, may explain the reduced turnout; voters with relatively moderate views may be less likely to turnout when both candidates adopt more extreme positions.

Of course, an important caveat is worth noting. Under the regression discontinuity approach, I am – by necessity – focusing on candidates who barely made it through their primary. This may indicate that they are weaker candidates or have been weakened by a difficult primary. Voters' responses to candidate gender amongst more competitive candidates may differ. Also, House elections may receive less elections than Senate, gubernatorial, or presidential elections; it is possible that increased attention and scrutiny may exacerbate any effects of candidate gender. Recall, however, that Wolak's (2014) experimental study finds that female participants are unresponsive to the gender of hypothetical candidates (holding other candidate characteristics constant), while male participates indicate they would be less likely to participate in an election with a female candidate. However, the fact that my results align relatively well with Wolak's (2014) hypothetical choice experiment suggests that the results may not be entirely dependent on identification strategy used in this paper.

References

[12] These results obviously also push back on the assertion that female candidates gain an advantage in the form of higher support from female voters.

Anastasopoulos, Lefteris. "Estimating the gender penalty in House of Representative elections using a regression discontinuity design." *Electoral Studies* (2016).

Anderson, Simon P., and Gerhard Glomm. "Alienation, indifference and the choice of ideological position." *Social Choice and Welfare* 9, no. 1 (1992): 17-31.

Anzia, Sarah F., and Christopher R. Berry. "The Jackie (and Jill) Robinson effect: why do congresswomen outperform congressmen?." *American Journal of Political Science* 55, no. 3 (2011): 478-493.

Besley, Timothy, and Stephen Coate. "An economic model of representative democracy." *The Quarterly Journal of Economics* (1997): 85-114.

Bhalotra, Sonia, and Irma Clots-Figueras. "Health and the political agency of women." *American Economic Journal: Economic Policy* 6, no. 2 (2014): 164-197.

Black, Jerome H., and Lynda Erickson. "Women candidates and voter bias: do women politicians need to be better?." *Electoral Studies* 22, no. 1 (2003): 81-100.

Bonica, Adam. "Database on Ideology, Money in Politics, and Elections: Public version 1.0 [computer file]." (2013).

Bonica, Adam. "Mapping the ideological marketplace." *American Journal of Political Science* 58, no. 2 (2014): 367-386.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82, no. 6 (2014): 2295-2326.

Carruthers, Celeste K., and Marianne H. Wanamaker. "Municipal Housekeeping." *The Journal of Human Resources* 50 (2015): 4.

Casas-Arce, Pablo, and Albert Saiz. "Women and power: unpopular, unwilling, or held back?." *Journal of Political Economy* 123, no. 3 (2015): 641-669.

Clots-Figuerasa, Irma. "Are female leaders good for education?." *American Economic Journal: Applied Economics* 4, no. 1 (2012): 212-244.

De Paola, Maria, Vincenzo Scoppa, and Marco Alberto De Benedetto. "The impact of gender quotas on electoral participation: Evidence from Italian municipalities." *European Journal of Political Economy* 35 (2014): 141-157.

Dolan, Kathleen. "Symbolic mobilization? The impact of candidate sex in American elections." *American Politics Research* 34, no. 6 (2006): 687-704.

Downs, Anthony. "An economic theory of political action in a democracy." *The Journal of Political Economy* (1957): 135-150.

Ferreira, Fernando, and Joseph Gyourko. "Do political parties matter? Evidence from U.S. cities." *Quarterly Journal of Economics* 124, no. 1 (2009): 399-422.

Ferreira, Fernando, and Joseph Gyourko. "Does gender matter for political leadership? The case of US mayors." *Journal of Public Economics* 112 (2014): 24-39.

Imbens, Guido, and Karthik Kalyanaraman. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies* (2011): rdr043.

Lawless, Jennifer L. "Politics of presence? Congresswomen and symbolic representation." *Political Research Quarterly* 57, no. 1 (2004): 81-99.

Llavador, Humberto. "Electoral platforms, implemented policies, and abstention." *Social Choice and Welfare* 27, no. 1 (2006): 55-81.

Lott, John R. "How dramatically did women's suffrage change the size and scope of government?." *Journal of Political Economy* 107, no. 6 Part 1 (1999): 1163-1198.

McCrary, Justin. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142, no. 2 (2008): 698-714.

McElroy, Gail, and Michael Marsh. "Candidate gender and voter choice: Analysis from a multimember preferential voting system." *Political Research Quarterly* (2009).

Osborne, Martin J., and Al Slivinski. "A model of political competition with citizen-candidates." *The Quarterly Journal of Economics* (1996): 65-96.

Pettigrew, Stephen, Karen Owen, and Emily Wanless. (2014). "U.S. House Primary Election Results (1956-2010)",http://dx.doi.org/10.7910/DVN/26448, Harvard Dataverse, V3.

Reingold, Beth, and Jessica Harrell. "The Impact of Descriptive Representation on Women's Political Engagement: Does Party Matter?." *Political Research Quarterly* 63, no. 2 (2010): 280-294.

Wolak, Jennifer. "Candidate Gender and the Political Engagement of Women and Men." *American Politics Research* 43, no. 5 (2015): 872-896.

Tables

Table 1: Summary statistics

	(1) Means: Full sample (All elections where one party's primary was between a man and a woman)	(2) Means: Subsample: Within optimal bandwidth (Primary margin of victory < 23%)	(3) Means: Subsample: Within half opt. bandwidth (Primary margin of victory < 12%)
Outcomes			
Randomly selected candidate's general election vote share	0.49	0.44	0.44
	(0.01)	(0.01)	(0.01)
Turnout: Votes per capita	0.32	0.32	0.32
	(0.00)	(0.00)	(0.01)
Turnout: Total votes cast	197,021.42	198,579.15	197,077.64
	(1,911.04)	(2,874.81)	(3,612.90)
Candidate characteristics			
Randomly selected candidate's party (Democrat=1, Republican=0)	0.63	0.56	0.54
	(0.02)	(0.02)	(0.03)
Randomly selected cand. incumb. status (Incumb.=1, Challenger=0)	0.27	0.05	0.04
	(0.01)	(0.01)	(0.01)
Randomly selected state leg. exp. (Any state leg. exp.=1, 0 otherwise)	0.25	0.20	0.18
	(0.01)	(0.02)	(0.02)
Opposing candidate's incumb. status (Incumb.=1, Challenger=0)	0.48	0.58	0.58
	(0.02)	(0.02)	(0.03)
Geographic characteristics			
Region: Northeast	0.17	0.17	0.18
	(0.01)	(0.02)	(0.02)
Region: Midwest	0.27	0.26	0.23
	(0.01)	(0.02)	(0.03)
Region: South	0.27	0.26	0.28
	(0.01)	(0.02)	(0.03)
Region: West	0.30	0.30	0.30
	(0.01)	(0.02)	(0.03)
Congressional district pop.	623,143.45	624,106.57	622,671.18
	(3,508.11)	(5,355.34)	(6,720.06)
Observations	957	420	243

Notes: This table draws on the CQ Election Returns data to summarize characteristics of the elections used in the estimating sample. Column 1 includes the full sample of elections between 1972-2010 for which one party's primary was between a male and female candidate. Columns 2 and 3 restrict the sample to observations where the mixed-gender primary was decided by a narrower margin of victory as indicated (the optimal bandwidth selected in accordance with the Calonico et al. procedure). Standard errors are in parentheses.

Table 2: RD Estimates: Impact of candidate gender on the election outcome and turnout, drawing on election returns data

VARIABLES	(1) Randomly selected candidate's vote share	(2) Randomly selected candidate's vote share	(3) Turnout: Votes per capita	(4) Turnout: Votes per capita	(5) Turnout: Ln(Total votes)	(6) Turnout: Ln(Total votes)
Panel A: Main results						
Female wins	0.002	0.001	-0.020*	-0.019*	-0.082**	-0.077*
	(0.030)	(0.024)	(0.011)	(0.011)	(0.039)	(0.040)
Add'l controls		X		X		X
Observations	420	420	420	420	420	420
R-squared	0.091	0.466	0.438	0.448	0.489	0.504
Panel B: Allowing for heterogeneity by party of randomly assigned candidate						
Fem. Wins X Dem	0.004	-0.005	-0.031*	-0.030*	-0.110*	-0.104*
	(0.043)	(0.034)	(0.017)	(0.017)	(0.058)	(0.058)
Fem. Wins X Rep	-0.007	0.005	-0.005	-0.004	-0.043	-0.041
	(0.033)	(0.026)	(0.017)	(0.017)	(0.055)	(0.057)
Add'l controls		X		X		X
Observations	420	420	420	420	420	420
R-squared	0.111	0.467	0.443	0.453	0.492	0.509

Notes: All specifications in this table are local linear regressions, restricted to a bandwidth of roughly 0.23 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies. Columns 5 and 6 control for log of total population within the Congressional district.

Additional controls (where indicated): Incumbency status of randomly selected candidate, a dummy variable indicating prior state legislative experience of randomly selected candidate, incumbency status of opposing candidate.

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3: ANES Summary Statistics

Variable	Mean (Std. error)	Variable	Mean (Std. error)
Voted: House election	0.58	Race: White	0.79
(1=yes, 0=no)	(0.01)		(0.01)
Party: Dem.	0.38	Race: Black	0.12
	(0.01)		(0.01)
Party: Repub.	0.27	Race: Asian	0.01
	(0.01)		(0.00)
Party: Indep.	0.27	Race: Nat. American	0.01
	(0.01)		(0.00)
Party: Other	0.01	Race: Hispanic	0.06
	(0.00)		(0.00)
Party: No pref.	0.07	Race: Other	0.00
	(0.00)		(0.00)
Age	45.79	Female	0.56
	(0.30)		(0.01)
Observations		3,459	

Table 4: RD Estimates: Impact of candidate gender on ANES respondents' likelihood of voting in the relevant Congressional election

VARIABLES	(1) Voted	(3) Voted	(3) Voted
Female wins	-0.091		
	(0.070)		
Fem. wins X Repub. cand.		0.063	
		(0.090)	
Fem. wins X Dem. cand.		-0.296***	
		(0.063)	
Fem. wins X Female voter			0.030
			(0.069)
Fem. wins X Male voter			-0.231**
			(0.094)
Observations	1,113	1,113	1,113
R-squared	0.149	0.159	0.154

Notes: The outcome variable is equal to one if the respondent reports voting in the Congressional election in their district. All specifications in this table are local linear regressions, restricted to a bandwidth of roughly 0.19 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies and additional controls (Party affiliation of respondent, party affiliation of randomly selected candidate, interaction of party of respondent and candidate, gender, race, and age).

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: RD Estimates: Impact of candidate gender on respondents' likelihood of voting in the relevant Congressional election – testing for heterogeneity by voter-candidate party congruence

VARIABLES	(1) Voted	(2) Voted: Subsample: Randomly selected cand. is Republican	(3) Voted: Subsample: Randomly selected cand. is Democrat	(4) Voted: Subsample: Male voters	(5) Voted: Subsample: Female voters
Fem. wins X (Voter Party = Cand. Party)	-0.096	0.094	-0.321***	-0.290**	-0.038
	(0.085)	(0.113)	(0.094)	(0.114)	(0.104)
Fem. wins X (Voter Party ≠ Cand. Party)	-0.088	0.070	-0.377**	-0.399***	0.082
	(0.092)	(0.140)	(0.151)	(0.110)	(0.098)
Observations	1,113	472	641	512	601
R-squared	0.150	0.189	0.151	0.230	0.172

Notes: The outcome variable is equal to one if the respondent reports voting in the Congressional election in their district. All specifications in this table are local linear regressions, restricted to a bandwidth of roughly 0.19 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies and additional controls (Party affiliation of respondent, party affiliation of randomly selected candidate, interaction of party of respondent and candidate, gender, race, and age).

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: RD Estimates: Impact of candidate gender on candidates' ideological positioning and divergence

VARIABLES	(1) Republican candidate's Dynamic CF-Score	(2) Democrat candidate's Dynamic CF-Score	(3) Divergence in CF-Score (Dem. Score – Repub. Score)
Mean of outcome variable	*0.911*	*-0.796*	*1.710*
(Standard dev.)	*(0.302)*	*(0.470)*	*(0.529)*
Panel A: Randomly assigned candidate is Republican			
Female wins (for Repubs)	-0.039	-0.004	-0.035
	(0.065)	(0.091)	(0.099)
Observations	220	220	220
R-squared	0.206	0.191	0.172
Panel B: Randomly assigned candidate is Democrat			
Female wins (for Dems)	0.147*	-0.208*	0.346***
	(0.077)	(0.116)	(0.123)
Observations	237	237	237
R-squared	0.178	0.242	0.243

Notes: All specifications in this table are local linear regressions, restricted to a bandwidth of roughly 0.31 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies, incumbency status of randomly selected candidate, and incumbency status of opposing candidate.

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figures

Figure 1: Binned scatterplots: Narrow primary victory of female candidates and general election outcomes

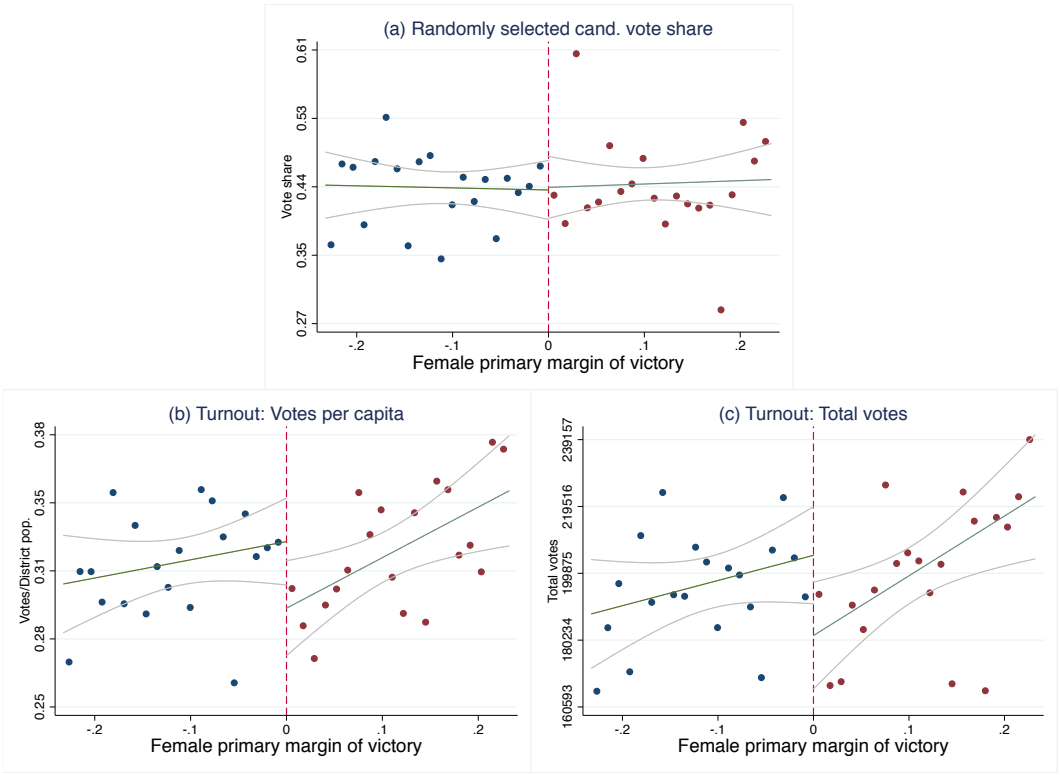

Figure notes: "Female primary margin of victory"=[Female candidate's primary vote share – Male candidate's primary vote share]. Female primary margin>0 implies a female candidate won the primary and represents her party in the general election. Outcome variables (vote share and turnout) refer the general election that the randomly assigned candidate participates in. Fitted lines are simple linear with no controls. Points are binned averages of raw data.

APPENDIX: Additional tables & figure

Table A1: RD Validity: Covariate balance tests

VARIABLES	(1) Lagged vote share for Democrat candidate	(2) Lagged vote share for randomly selected candidate's party	(3) Randomly selected candidate's party (Democrat=1)	(4) Opposing candidate's incumbency status (Incum.=1)	(5) Randomly selected candidate's incumbency status (Incum.=1)	(6) Randomly selected candidate's prior state leg. exp (Any exper.=1)
Female wins	0.009	0.022	0.097	-0.033	0.019	-0.109
	(0.039)	(0.033)	(0.099)	(0.085)	(0.032)	(0.073)
Observations	317	310	420	420	420	420
R-squared	0.076	0.040	0.072	0.075	0.064	0.040

Notes: All specifications in this table are local linear regressions, restricted to a bandwidth of roughly 0.18 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies.

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A2: RD Estimates: Impact of candidate gender on the election outcome and turnout, drawing on election returns data (Half and double the optimal bandwidth)

VARIABLES	(1) Randomly selected candidate's vote share	(2) Randomly selected candidate's vote share	(3) Turnout: Votes per capita	(4) Turnout: Votes per capita	(5) Turnout: Ln(Total votes)	(6) Turnout: Ln(Total votes)
Panel A: Half optimal bandwidth around primary election cutoff						
Female candidate	-0.016	0.003	-0.029*	-0.028*	-0.120**	-0.126**
	(0.041)	(0.031)	(0.015)	(0.015)	(0.055)	(0.053)
Add'l controls		X		X		X
Observations	243	243	243	243	243	243
R-squared	0.154	0.485	0.445	0.448	0.460	0.469
Panel B: Double optimal bandwidth around primary election cutoff						
Female candidate	0.002	0.013	-0.009	-0.010	-0.038	-0.044
	(0.019)	(0.018)	(0.009)	(0.009)	(0.033)	(0.034)
Add'l controls		X		X		X
Observations	661	661	661	661	661	661
R-squared	0.071	0.495	0.383	0.395	0.478	0.495

Notes: All specifications in this table are local linear regressions, restricted to either half or double the optimal bandwidth of roughly 0.18 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies. Columns 5 and 6 control for log of total population within the Congressional district.

Additional controls (where indicated): Incumbency status of randomly selected candidate, a dummy variable indicating prior state legislative experience of randomly selected candidate, incumbency status of opposing candidate.

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A3: RD Estimates: Impact of candidate gender on the election outcome and turnout, drawing on election returns data (Global parametric RD approach)

VARIABLES	(1) Randomly selected candidate's vote share	(2) Randomly selected candidate's vote share	(3) Turnout: Votes per capita	(4) Turnout: Votes per capita	(5) Turnout: Ln(Total votes)	(6) Turnout: Ln(Total votes)
Panel A: 2nd Degree Polynomial in primary election margin of victory						
Female wins	0.005	0.007	-0.010	-0.010	-0.050	-0.050
	(0.022)	(0.018)	(0.009)	(0.010)	(0.034)	(0.035)
Add'l controls		X		X		X
Observations	957	957	957	957	957	957
R-squared	0.235	0.592	0.354	0.368	0.477	0.493
Panel B: 3rd Degree Polynomial in primary election margin of victory						
Female wins	0.004	0.006	-0.017	-0.017	-0.074	-0.075
	(0.033)	(0.028)	(0.015)	(0.015)	(0.055)	(0.055)
Add'l controls		X		X		X
Observations	957	957	957	957	957	957
R-squared	0.236	0.593	0.355	0.369	0.478	0.494
Panel C: 4th Degree Polynomial in primary election margin of victory						
Female wins	-0.009	-0.001	-0.017	-0.018	-0.085	-0.087
	(0.037)	(0.027)	(0.016)	(0.015)	(0.058)	(0.058)
Add'l controls		X		X		X
Observations	957	957	957	957	957	957
R-squared	0.239	0.595	0.355	0.370	0.478	0.495

Notes: All specifications in this table are global polynomial regression discontinuity estimates. Each panel fits "female margin of victory" using a different degree of polynomial (2nd, 3rd, or 4th). I report only the coefficient on "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. Other coefficients are suppressed for space. All specifications include year dummies. Columns 5 and 6 control for log of total population within the Congressional district.

Additional controls (where indicated): Incumbency status of randomly selected candidate, a dummy variable indicating prior state legislative experience of randomly selected candidate, incumbency status of opposing candidate.

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A4: RD Estimates: Impact of candidate gender on ANES respondents' likelihood of voting in the relevant Congressional election (Alternative specifications)

VARIABLES	(1) Voted	(2) Voted	(3) Voted	(4) Voted	(5) Voted	(6) Voted
Female wins	-0.165	-0.091	-0.223	-0.175*	-0.086	-0.043
	(0.111)	(0.070)	(0.134)	(0.093)	(0.075)	(0.053)
Add'l controls		X		X		X
Bandwidth	Optimal (19%)	Optimal (19%)	1/2 X Optimal (9.5%)	1/2 X Optimal (9.5%)	2 X Optimal (38%)	2 X Optimal (38%)
Observations	1,130	1,113	596	585	1,902	1,869
R-squared	0.056	0.149	0.082	0.170	0.050	0.149

Notes: All specifications in this table are local linear regressions, restricted to specified bandwidths around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies.

Additional controls (where indicated): Party affiliation of respondent, party affiliation of randomly selected candidate, interaction of party of respondent and candidate, gender, race, and age.

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A5: RD Estimates: Impact of candidate gender on ANES respondents' likelihood of reporting positive feelings about randomly assigned candidate

VARIABLES	(1) Respondent reports positive feelings about rand. assigned cand. (Yes=1)	(2) Respondent reports positive feelings about rand. assigned cand. (Yes=1)	(3) Respondent reports positive feelings about rand. assigned cand. (Yes=1)	(4) Respondent reports positive feelings about rand. assigned cand. (Yes=1)
Female wins	0.182			
	(0.113)			
Fem. wins X Repub. cand.		-0.027		
		(0.163)		
Fem. wins X Dem. cand.		0.202*		
		(0.114)		
Fem. wins X (Voter Party = Cand. Party)			0.215*	
			(0.114)	
Fem. wins X (Voter Party ≠ Cand. Party)			0.146	
			(0.146)	
Fem. wins X Fem. voter				0.204
				(0.133)
Fem. wins X Male voter				0.166
				(0.131)
Observations	630	630	630	630
R-squared	0.089	0.101	0.090	0.092

Notes: The outcome variable is equal to one if the respondent reports positive feelings about the randomly assigned candidate (the candidate who narrowly emerged from a mixed-gender primary). All specifications in this table are local linear regressions, restricted to a bandwidth of roughly 0.19 around the cutoff (primary election margin of victory=0). Coefficients on "female margin of victory" and "female wins * female margin of victory" are suppressed. I report only the main effect of "Female wins," which is interpreted as the causal effect of a female candidate serving as a candidate in the general election. All specifications include year dummies and additional controls (Party affiliation of respondent, party affiliation of randomly selected candidate, interaction of party of respondent and candidate, gender, race, and age).

Robust standard errors (clustered at state-decade level) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figure A1: RD Validity check: Distribution of female margin of victory in primary election

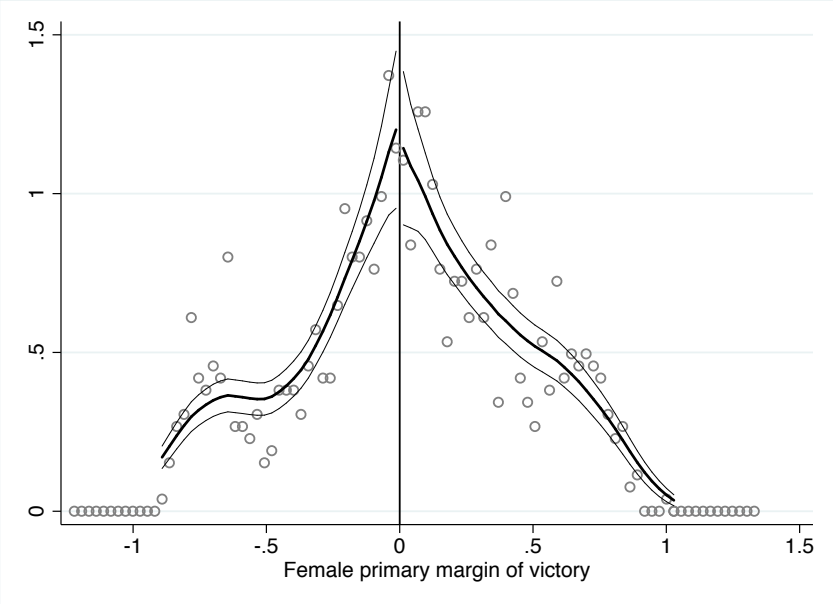

Figure notes: "Female primary margin of victory"=[Female candidate's primary vote share – Male candidate's primary vote share]. Female primary margin>0 implies a female candidate won the primary and represents her party in the general election. Code to generate the graph drawn from McCrary (2008).

Figure A2: RD Validity check: Binned scatterplots: Narrow primary victory of female candidates and other covariates

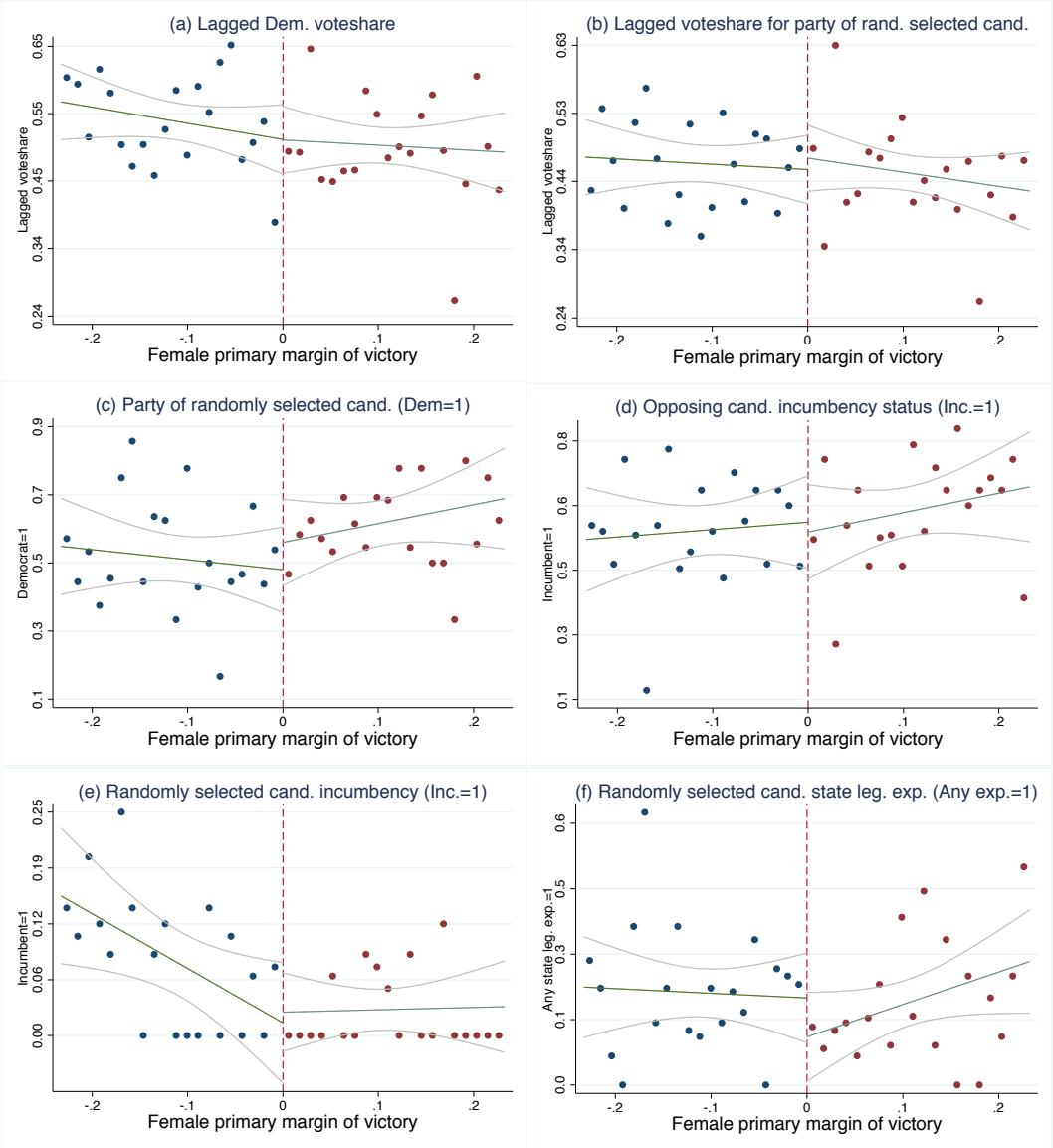

Figure notes: "Female primary margin of victory"=[Female candidate's primary vote share – Male candidate's primary vote share]. Female primary margin>0 implies a female candidate won the primary and represents her party in the general election. Outcome variables (vote share and turnout) refer the general election that the randomly assigned candidate participates in. Fitted lines are simple linear with no controls. Points are binned averages of raw data.