The Double-edged Effect of Political Representation

Dayea Oh∗ †

August 2022

Abstract

I analyze measures of political ideology and results from U.S. House of Representatives elections spanning the period 1956 through 2014 to show the impact of woman running for office on the competing candidate ideology. Using a regression discontinuity design combined with a bounding strategy, I find that legislators who run against women show no change in ideology on the liberal-conservative scale but vote more conservatively on the cultural and racial issues in Congress, conditional on winning the race. This effect only exists for the Democratic legislators and translates to a 0.5 to 0.7 standard deviation shift towards the party extreme in their ideological score. The result supports Calvert-Wittman candidate location model with ideologically motivated candidates diverging from the center to align closer to party ideal when faced with opponents further away from the center.

∗Harvard Kennedy School, Cambridge MA 02139. ohd@g.harvard.edu
†I am grateful to seminar participants at Harvard University, Michela Carlana, Will Dobbie, Reshmaan Hussam, Vincent Pons, James Snyder, and Ebonya Washington for their valuable comments.
I. Introduction

In March 2018, New York City saw an unlikely candidate to challenge a long-standing incumbent governor Andrew Cuomo. Cynthia Nixon, an actress known for her role as Miranda Hobbes in the HBO’s comedy Sex in the City, announced her candidacy against Cuomo. Although Nixon eventually lost the primary election with receiving 34% of the vote, many experts believed that it was her campaign against Cuomo that pushed him leftward, including changing his stance on legalizing recreational marijuana, putting a ban on a single-use plastic bags, as well as vowing to restore the voting rights of the felons (The Atlantic 2018). Although just one of the many elections in the U.S., this 2018 New York gubernatorial election exemplifies how a candidate, even with an evidently low probability of winning, can exert influence on the other candidates’ political actions by taking part in the race and addressing issues otherwise would have been left under the table.

This paper examines the impact of women running for office on competing candidate ideology. Can an opponent, simply by participating in the race, exert influence on the winning candidate? If candidates adjust their platform and campaign strategy to their opponent’s identity and platform, this may translate to changes in legislative behavior later in Congress. While the idea that the candidates are influenced by their opponent seems apparent, empirical evidence to causally support this notion is surprisingly scarce. I tackle this question in the setting of U.S. House of Representatives elections spanning the period 1956 through 2014 using candidate political ideology scores measured before and after the general election. Specifically, I focus on the impact of one particular trait of opponent – gender – on candidate ideology.

Providing a credible answer to this empirical question has proven difficult for a couple of reasons. First, the opponent’s decision to run for office is endogenous to candidate ideology. I resolve this issue by exploiting a quasi-experiment embedded in the U.S. Congressional electoral system that generates essentially random variation in opponent gender. I study the set of electoral races in which the opponent had barely won the primary against an opposite-sexed runner-up. The key identifying assumption is that the candidates whose female opponent barely won the primary race against a male runner-up are comparable to...
candidates whose male opponent barely won the primary race against a female runner-up. This regression discontinuity design then addresses the endogeneity problem by isolating arguably exogenous variation in opponent gender.

But even if the opponent’s decision to enter the race is exogenous, the candidate winning against the opponent is not. If the candidate ideology is measured only after the election and we do find an impact, two separate but indistinguishable potential channels may be at work: (i) the candidate running against the opponent may shift candidate ideology and (ii) the candidate winning against the opponent may change the final pool of winning candidates with already different ideologies at baseline. Then the second empirical challenge lies in isolating the effect driven by the first channel. I surmount this issue by showing that the opponent gender does not affect a candidate’s probability of winning the race, and thereby claiming that if an effect exists, it is driven by the first channel only. I further supplement this argument by relaxing the assumption that the opponent gender has no effect on the candidate’s probability of winning. This renders comparison of winners with close female versus male opponents biased, as the type of candidates who win against female and male would be different. Following the identification strategy in Anagol and Fujiwara (2016) and Granzier, Pons, and Tricaud (2021), I estimate the effect on candidate ideology conditional on winning the race by making assumptions about the unobserved ideology of the losers using the potential outcomes framework and finally provide bounds on the potential treatment effect.

My first result establishes a null effect of the gender of the opponent on candidate ideology during the electoral campaign stage. Using the 1980-2014 Congressional election results linked with Database on Ideology, Money in Politics, and Elections (Bonica 2016) as a measure of political ideology of the candidates, I find that running against a female opponent has no significant impact on the candidate ideology before the election compared to running against a male opponent. The absence of the ideological change does not arise from the stationary nature of the ideology scores. The presence of an opponent in the general election and the candidate vote share both show strong association with candidate ideology, suggesting the candidates do respond to the competition in the race but that the change is simply not driven by the gender of the opponent, at least before the election and on the
conservative-liberal scale.

The second result shows that the gender of the opponent can leave a long-term impact on candidate ideology on a racial and cultural dimension that lasts beyond the election. I put together the primary and general election results from 1956–2014 Congressional races and assign the winners of the race their DW-Nominate scores (Poole and Rosenthal 1985, 2007; Nokken and Poole 2004) to measure the underlying legislator ideology in Congress. Consistent with the short-term impact, the gender of the opponent has no effect on the legislators’ voting pattern on the liberal-conservative scale. Instead, I find a significant shift in the cultural and racial dimension of ideology for the Democratic legislators. Democratic legislators who run against female GOP opponents move 0.5 to 0.7 standard deviation away from the center towards the party extreme to be more racially and culturally conservative.

This work offers two primary contributions. First, I speak to the literature on candidate positioning under electoral competition. There are two standard views on candidate positioning. First is the most famous and widely accepted theory in the literature of electoral competition, Downs (1957)’ median voter theory. Downsian political candidates are described as “politically motivated” or “office motivated” and their sole objective is to win the elections, where they compete in a single dimensional policy space. This pressure of electoral competition constrains the candidates’ policies to the one most preferred by the median voter, giving way to the crucial implication of the model that in a two-party system, full convergence of policies result in equilibrium. The outcome still holds when we allow candidates to be ideologically motivated differently. Wittman (1977, 1983) and Calvert (1985) write that with uncertainty about the preference of the median voter, these “ideologically motivated” politicians may not fully converge on an identical platform but would still compromise on a more moderate point relative to their respective most preferred policies, resulting in partial convergence.

The second alternative view on policy choice formally accounts for the policy divergence when the assumption of credible commitment to platform is relaxed. In the citizen-candidate model of Osborne and Slivinski (1996) and Besley and Coate (1997), candidates are allowed different underlying policy preferences but without a precommitment device to the announced policies. The only credible outcome then is for the candidates to adopt their
most preferred policy in the office.¹ That is, electoral pressure does not compel opposing candidates to adjust their positions as Downs would predict; regardless of the identity or the ideology of the opponent, the only time-consistent equilibrium in the citizen-candidate model is for each candidate to carry out her ex post most preferred policy when in office.

Then the two models present us with two different predictions of how a candidate may respond to an opponent. Downsian and the Calvert-Wittman model would tell us that the candidates would respond to the shift in the preferences of the voters with the goal of winning the race and therefore adjust their positions as a response to the opponent. And while the citizen-candidate model predicts preferences and personal traits of the candidate to influence their political stance and policy making,² preferences and personal trait of the opponent should not affect the policies the candidate implement in the office. After their decision to enter the race has been made, the only credible policy choice is to implement their most preferred policy.

My paper adds to the literature by showing for the first time the effect of an opponent on candidate ideology. Candidates running in contested races tend to lean more extreme compared to when running unopposed, and the effect size increases with the level of electoral competition.³ But perhaps more importantly, it is not only the threat of competition that affects candidate ideology but also the identity of the opponent that can shift legislator voting behavior even after the election. I find that Democratic legislators vote more conservatively on racial and cultural issues in Congress if they had run against a female GOP opponent.

¹ More specifically, Alesina (1988) demonstrates that the ability to credibly commit is an essential assumption of the Downsian model. He shows that in a one shot game, the only time-consistent equilibrium is the one where candidates carry out their ex post most preferred policy, a full divergence of the policies: If the candidates are ideologically motivated and there’s no possibility of binding precommitments to the announced policies, rational and forward-looking voters will take this into account, rendering the compromised centrist promises noncredible and leading each candidate to announce and pursue his most preferred policy.

² Empirical papers that support the citizen-candidate model include Washington (2008), Chattopadhyay and Dubois (2004), Clots-Figueras (2011), and Ferreira and Gyourko (2014).

³ The effect of electoral competition on legislator ideology has been studied before with mixed results. Lee, Moretti, and Butler (2004) find that the degree of electoral strength has no effect on a legislator’s voting behavior. The authors use the data on Americans for Democratic Action (ADA) scores for all Representatives in the U.S. House from 1946-1995 linked to the election returns data, where the ADA scores choose about twenty high-profile roll-call votes to create an index for each Representative of the House. They find similar results with alternative measures of voting records including the DW-Nominate scores. Although our papers differ in sample, time period, and the measure of ideology, perhaps the most prominent difference rises from our samples where my sample includes both the winners and the losers of the race while the authors only uses the winners of the race (legislators). Finally, Gerber and Morton (1998) find that state legislators elected under closed primaries take more extreme positions than legislators elected under open primaries.
ideological shift of candidates as a response to the opponent is distinctively Downsian as the model predicts the politicians to change their stance depending on the electoral competition, adopting their policies and issue space to their opponent and consequently to the voter preferences. The direction of the result is also consistent with the predictions of the extension of the Downsian spatial model where the candidates compete on non-policy dimensions as well as policy issues. If one candidate possess known quality advantage that gives her relative electoral strength, the other candidates strategically adopt more divergent and extreme platforms to differentiate themselves on the policy dimension (Ansolabehere and Snyder 2000; Groseclose 2001; Aragones and Palfrey 2002).

Second, this paper speaks to and expands the existing literature on the effect of women in office. Previous literature has mostly focused on the impact of women already in office. Studies show that women in political office affect how and where public funds are spent, and increase participation of future generation of female politicians. Female representatives introduce and pass more priority bills dealing with issues of women, children, and families more than men (Thomas 1991; Swers 1998), invest more in public goods that directly help women (Chattopadhyay and Duflo 2004), and improve perception of female leader effectiveness and weaken stereotypes about gender roles to bring future electoral gains for women (Beaman et al. 2009). They also serve as role models and increase participation of future generation of female politicians. Presence of high profile female politicians inspire interest of political activism among young women (Campbell and Wolbrecht 2006), inspire other women to run for office (Ladam, Harden, and Windett 2018), help female council candidates advance more from their initial list rank (Baskaran and Hessami 2018), and even foster growth in education attainment of girls (Beaman et al. 2012).

But many more women in fact run for office. In 2018, 519 women ran in the Congressional election while 116 won. In 2020, 643 women ran in the Congressional election while 126 won. For gubernatorial, state executive, state legislative, and local elections, the numbers look similar. And while we understand the many ways the 126 Congresswomen can yield their influence on the policies, public spending, and citizens, we have little understanding of the effect the rest of 643 women have. To the best of my knowledge, this paper is the first work that identifies the impact of women running in Congressional races on the competing
candidates, before and after the election.\footnote{Dolan (2008) studies whether female opponents shape male candidate behaviors by comparing campaign websites of men who ran against women versus men who ran against other men to find no evidence that men “play” to women voters by investing more in women’s issue, but suffers from endogeneity issues. Meyer(2018) investigates the effect of opponent gender on Congressional candidates’ issue emphasis on social media to find that male candidates decrease their attention to women’s issues when running against a male opponent versus female. Neither paper includes the effect on the candidate after the election.}

The remainder of the paper is organized as follows. In Section II, I describe the data on political ideology and election results. In Section III, I examine the short term effect of electoral competition on candidate ideology, and next discuss the empirical strategy of the main analysis in Section IV. Section V provides the results, and Section VI concludes.

\section*{II. Data}

In this section, I describe the sources of my data and present summary statistics.

\subsection*{II.A. U.S. House election results}

I collect the U.S. House general election results for 1956–1974 from the Database of Historical Congressional Statistics (Swift et al. 2009) and the results for 1976–2014 general elections from MIT Election Data and Science Lab (2017). All include information on the year the election was held, the state, Congressional district, party, candidate names, and the number of votes each candidate received in the race. I limit my sample to the major parties (GOP and Democratic) candidates, and drop special elections as well as elections held in Louisiana, California from 2010, and Washington from 2006 due to their majority-vote system.\footnote{In majority-vote system, all candidates running for a local, state, or federal office appear on the same ballot regardless of their partisan affiliations. If a candidate wins a simple majority of all votes (i.e., 50 percent, plus one vote), he or she wins the election outright. If no candidate meets that threshold, the top two finishers, regardless of their partisan affiliations, advance to a second – considered the general in other states – election. In the second round, the candidate who receives the greatest number of votes wins (Balletopia, 2021).} Candidates who ran in the elections but withdrew or resigned from the post, promoted to a higher office, or died after winning the general election but before starting his/her term are also dropped.

I merge the above general elections data to the primary elections data from Pettigrew,
The primary elections data has information on the races including the year, state, district, party, candidate name and the number of votes each candidate received as well as information on the incumbency status and the gender of the candidate. The primary runoff elections data is unavailable for years before 1994 and as a result 73 general election candidate × year pair is left unmatched to the primary results. They amount to 0.3% of the whole sample. Finally, 100 cases were also left unmatched with the primary election results due to circumstances such as candidates withdrawing or resigning from the race, being promoted to higher office, or dying after the primary election but before the general election. The winners of the primary in these cases also do not match the replacing candidates who end up running in the general election in stead and the merge is left unsuccessful. They amount to 0.4% of all general election candidate × year pair and are dropped from the sample.

My final sample consists of 23,271 candidate × year pairs in the U.S. House of Representatives general elections between 1956–2014, of which 2,197 are female. Out of the 2,197 cases, 1,001 ended with the female candidate winning the race. Figure I shows the time trend of the fraction of the female candidates who ran in the U.S. House of Representatives general election. The number gradually rises with a sharp increase in 1992, the year dubbed as the “Year of the Woman” where a record-breaking four women were elected as senators and 24 women as representatives to Congress.

II.B. Measures of political ideology

I now discuss the choice of the dependent variable. There are several ways to measure the political ideologies of the elected officials and sometimes of the unelected nonincumbent candidates. They include constructing scores using legislature roll call votes (Poole 2011; Shor and McCarty 2011; Steven D. Levitt and James M. Snyder 1999), surveying the candidates (Ansolabehere, Snyder, and Stewart 2001; Burden 2004; Shor and Rogowski 2018), analyzing followers on Twitter (Barberá 2014), perceptions of civilians or experts...
Aldrich and McKelvey 1977; Hare et al. 2014; Ramey 2016), or the campaign contributions received by the candidates (Bonica 2013, 2014).

I focus on two measures in this paper, for their wide adoption in the literature: campaign finance scores to measure the impact on candidate ideology before the election, reflecting electoral politics, and the DW-Nominate scores to measure the impact after the election, reflecting congressional politics. The campaign finance (CF) scores are collected from the Database on Ideology, Money in Politics, and Elections (DIME) (Bonica 2016) and measure the political orientation of the candidates using the financial donations they receive during the electoral cycle. For each election held in 1980–2014, DIME tabulates campaign donations that’s greater than $200 using the mandatory reports required by the Federal Electoral Commission. With the assumption that the donors choose contributions to maximize the difference between the benefit from giving to particular candidates and the loss from giving to another whose ideological positions differ from their own, DIME provides an ideal point for each campaign donor and the candidate that encapsulates their ideology. This setting gives the CF score its great advantage that the data exists for both the winners and the losers of the election and we are able to measure the political orientation of both.

Although the basic versions of the CF scores are time-invariant, I use the time-varying CF scores for my analysis, which allows the candidates to have different scores in each electoral cycle. There are 13,924 U.S. House of Representatives general election candidate × year pair between 1980 and 2014 in my sample and 93% are matched with the corresponding CF scores. Figure II plots the time trend of the CF scores and Figure III shows the histogram of the scores averaged across years by party. As can be seen in Figures II and III, most CF scores lie between –2 and 2 and the higher scores correspond to more “conservative” ideologies. A few well known former Congressman and Congresswomen with their non-dynamic CF scores include Senator Bernie Sanders with a score of –1.445, speaker of the House of Representatives Nancy Pelosi with a score of –1.077, former Vice President Mike Pence with a score of 1.212, and former United States Secretary of State Mike Pompeo with

---

7. The high construct validity of these scores is well illustrated in Autor et al. (2020).
8. Some candidates do not have CF score because they do not meet the requirements for inclusion in the scaling. Bonica (2016) requires that candidates raise funds from at least 25 unique contributors during a given election cycle to be included in the period-specific CF scores.
I use the DW-Nominate scores from Lewis et al. (2021) developed by political scientists Poole and Rosenthal (1997, 2007) as a measure of legislator ideologies in Congress and link them to the election results. The Nominate score is a seminal, powerful tool to illuminate a lot of American political history on relationships of roll calls, public policies, and issues. They use the legislators’ roll call vote choices of “yeas” and the “nays” for each legislation piece and scales the legislators by their ideological location in the issue space assuming that the legislators are maximizing their utility function. These scores can then provide a comprehensive point estimate of congressman’s political orientation in each and every Congress and categorize elected officials on a scale from liberal to conservative.

The relative strengths of the DW-Nominate scores as a voting records index are that they allow a consistent measure of comparison of any legislative behavior or congressional activities across time, include all roll-call votes instead of a subset of votes, and ignore the Representative’s political party as well as the legislative issue in question, making it arguably more comprehensive and exogenous than some other alternative indices. As a result, they have been widely used by political scientists and economist alike to examine questions such as relationships of political polarization and wealth disparity (McCarty, Poole, and Rosenthal 2003), whether voters affect or elect policies (Lee, Moretti, and Butler 2004), how the partisan composition of an electorate impacts the policies adopted by an elected representative (Jones and Walsh 2018), or whether rising import competition contributed to the polarization of U.S. politics (Autor et al. 2020), among many others.

The DW-Nominate scores capture the Congressional politics in at most two issue spaces. The scores in the first dimension quantifies the typical left-right (liberal-conservative) spectrum on the economic matters while scores in the second dimension represent attitudes on salient cultural and lifestyle issues of the day. Scores closer to 1 are described as conservative whereas scores closer to −1 are described as liberal. Score close to zero is described as moderate. The version of the DW-Nominate scores I use allow legislators’ ideal points to evolve flexibly over time, and does not assume legislator ideology to be fixed over time in different Congresses (Nokken and Poole 2004). Table I presents summary statistics of the DW-Nominate scores in my sample. The mean of the first dimension DW-Nominate score
for the U.S. House of Representatives in the 85th – 114th (1956–2014) Congress is −0.028 with a standard deviation of 0.039. The mean of the second dimensional score, whose higher score corresponds to more culturally and racially more conservative tendencies, is 0.019 with a standard deviation of 0.419. Figure IV shows the time trend of the Nominate scores by party for the first dimension and the second dimension from 1956 to 2014 and Figure V presents the histogram of the Nominate scores.

III. Short-term impact on electoral ideology

I first assess the impact of women running for office on the competing candidates’ ideology using the campaign finance scores. Specifically, I compare the ideological scores of the general election candidates running for 1980–2014 U.S. House of Representatives office with female opponents to those with male opponents. Blindly comparing candidates with female opponents and ones without would cause other factors, such as political progressivity of a district, to affect the outcome variable. To overcome this endogeneity issue, I exploit close primary races to generate essentially random variation in the gender of the opponent the candidate faces in the general election. I run the following regression:

\[ Y_{it} = \beta_0 + \beta_1 D_{it} + \beta_2 X_{it} + \beta_3 D_{it}X_{it} + \epsilon_{it} \]  

where \( Y_{it} \) is the campaign finance score of candidate \( i \) who ran in election \( t \) and \( D_{it} \) is a binary indicator that equals 1 if candidate \( i \)'s opponent in election \( t \) is a female. \( X_{it} \) is the running variable that measures the difference between the vote share of the top two female candidate and the male candidate in the primary election of \( i \)'s opponent such that if \( X > 0 \), then \( D = 1 \) and if \( X < 0 \), then \( D = 0 \). In other words, if a woman wins the primary of the opposing party, candidate \( i \) will face a female opponent in the general election. If instead a man wins the primary of the opposing party, candidate \( i \) will face a male opponent in the general election.\(^9\)

Table II presents the estimates of the impact of running against female opponent versus

---

\(^9\) While it matters for the analysis how candidate \( i \)'s opponent fared in his/her primary, I take candidate \( i \)'s electoral strength in his/her own primary as given.
male in the general election on candidate ideology measured over the election cycle but before
the election. Republican candidates shift their CF scores upwards by 0.032 while Democratic
candidates decrease their scores by 0.010 when their opponent is women. Neither estimates
are significant. This insignificant and near zero impact for both parties imply that facing
a female opponent does not necessarily affect the candidates differently than facing a male
opponent before the election.

In trying to understand if this null impact is due to the unvarying nature of the ideological
scores, I run a simple model to see if the candidate ideology scores respond to competition
at all, by comparing candidates in contested and uncontested races and with varying levels
candidate vote share. This additional analysis is made possible by making use of the fact
that many candidates run more than once for the same office. During 1980–2014, 1,903
candidates ran for the U.S. House of Representative office twice or more. By comparing
the same candidate who runs in different election cycles, we are able to hold observed and
unobserved candidate characteristics constant. I run the following candidate fixed effects
regression:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \lambda_i + \delta_{st} + \epsilon_{it} \tag{2}$$

where $Y_{it}$ is the ideology score of candidate $i$ in a given election cycle $t$, $T_{it}$ equals 1 if
candidate $i$ has an opponent in $t$ and 0 if unopposed, $\lambda_i$ is candidate FE and $\delta_{st}$ is state-
year FE. This will ensure whatever political event or sentiment that induces the race to be
contested and the candidates to change their ideology at the state-year level is accounted
for. The coefficient of interest is $\beta_1$ which then tells us the change in the candidate ideology
score when the candidate runs opposed versus runs unopposed. I also rerun equation (2)
with $T_{it}$ as the vote share of the candidate $i$ in $t$ to estimate the impact of the degree of
competition on candidate ideology.

Table III presents the estimates of the impact of running in a contested versus uncontested
race and with different levels of competitiveness on candidate ideology measured during the
election campaign cycle. Columns 1 and 2 of Panel A demonstrate that having an opponent
is associated with a 0.005 higher ideology score across both parties, though insignificant,
relative to running unopposed. Motivated by literature on partisan differences, I run the same regression for the two major parties separately in columns 3–6 and find opposite-signed effects for the two parties. Democratic candidates appeal to more liberal donors in a contested race versus uncontested ($p < 0.01$) while Republican candidates appeal to more conservative donors ($p < 0.01$), suggesting the same candidates adopt more centrist positions in easy races and show more extremist tendencies in more challenging races. The effect persists after controlling for incumbency status of the candidate. The direction of the treatment effect stays consistent when we study the impact of candidate vote share on candidate ideology in columns 3 to 6 in Panel B where the Democratic candidates increase their CF scores by 0.434 with one percentage point increase in their candidate vote share ($p < 0.01$) and Republican candidates decrease their CF scores by 0.342 ($p < 0.01$).

Table III tells us that candidate ideologies measured by CF scores do respond to electoral competition, suggesting the lack of impact in Table II is not due to the stationary nature of CF scores but possibly due to the lack of response from the candidates. That the candidates adjust their campaign strategy and ideology depending on their electoral strength in the race may not be surprising as we expect an uncontested race to be different in many regards from a contested race and competitive races from landslide victories.\(^\text{10}\) Candidates naturally position themselves relative to their opponents, and are expected to strategically choose which issues to bring up and how to frame them to the voters.

What is surprising is that while the candidates respond to the existence and to the level of competition, they do not seem to respond differently to female versus male opponents even at a low-stake environment during the electoral campaigns, despite gendered differences in legislative voting behavior (Figure X) and studies that show voters infer candidate ideology

\(^\text{10}\) While not the main focus of this paper, the empirical evidence that quantifies the causal relationship between electoral strength and candidate ideology is scarce with mixed results. Fiorina (1973) finds that electorally stronger incumbents tend to be more moderate than electorally weaker incumbents and Ansolabehere, Snyder, and Stewart (2001) find that incumbents show more moderate tendencies compared to challengers using survey data on House candidates. On the other hand, Lee, Moretti, and Butler (2004) use close U.S. House elections to estimate the impact of electoral strength on subsequent representatives’ roll call voting records to find that electoral strength has no effect on legislator voting behavior. I find that the same candidates adopt idealistic stances that are 0.04–0.08 standard deviation more aligned with party extremes when in more electorally challenging races compared to “easy” races where they adopt more moderate/centrist stance, suggesting a result closer to Fiorina’s and Ansolabehere’s than to Lee’s. The difference may be a result of different time period and/or sample size.
on the basis of gender (Koch 2000, 2002; King and Matland 2003; Sanbonmatsu and Dolan 2009) and believe women to have competency over different political issues to men (Alexander and Andersen 1993; Huddy and Terkildsen 1993; Dolan 2010).

One possible reason for the lack of response from the candidates may be due to the high salience nature of Congressional elections. Because stereotypes function to fill in missing information, there may be no space for gender of the candidate to provide additional information in races where individual candidate’s political agenda and platform are clearly conveyed to voters and to the competing candidate. After all, while gender does send low informational cues on the ideological orientation of the candidates, political party is the primary determinant of congressional ideology and party stereotypes are more powerful than gender stereotypes (Hayes 2011).11

Another potential reason for the null effect arises from the unidimensional issues space of the CF scores. CF scores capture candidates’ ideology in the traditional conservative-liberalness spectrum only. This means that if the candidates do respond to the female opponents by changing their stances on “women” issues but if such issues are not correlated with the standard conservative-liberal scale in a linear manner, this change would not be captured by and result in changes in the CF scores. The women’s issue space may be operational on a different dimension other than the standard economical dimension, as studies have repeatedly shown that unidimensional model of ideology provides an incomplete basis for political ideology and minimum of two dimensions – economic and social ideology – are needed to account for domestic policy preferences (Poole and Rosenthal 1997; Feldman and Johnston 2014).

Despite the lack of impact, the effect may exist in a different policy dimension and/or in the long term after the candidates have won the race. In the remainder of the paper, I explore the long term impact of women running in the race on the legislative voting behavior of the opposing candidates after the general election.

11. Thus, an impact may exist at low information race such as local elections or primary races that lack the party cue to cleanly divide the ideological orientation of the running candidates.
IV. Long-term impact on congressional ideology

In this section, I discuss the empirical strategy for estimating the impact of opponent gender on candidate ideology after the election.

Empirical framework

I rely on the same regression discontinuity design to generate random variation in the opponent gender and to estimate the long term impact of facing women in the race:

$$ Y_{it} = \beta_0 + \beta_1 D_{it} + \beta_2 X_{it} + \beta_3 D_{it} X_{it} + \epsilon_{it} $$

(3)

where now $Y_{it}$ is the Nominate score of candidate $i$ who ran in race $t$, measuring candidate ideology after the election in Congress. All other variables are defined the same way as in equation (1). The key problem for inference is that although the above regression discontinuity design allows the opponent gender to be random, the candidate winning against the opponent is not, making our coefficient of interest $\beta_3$ a combined effect of two potential channels. First, the candidate ideology may shift as a result of running against the female opponent. This is the channel that we are interested in identifying. Second, we may find an impact because the candidates who tend to win against women may have different ideology to begin with, compared to the candidates who tend to lose against women.

I employ a two-parts analyses to isolate the effect that comes from the first channel only. In the first part of the analysis, I start by showing that the gender of the opponent in fact has no impact on the candidate probability of winning the race. This implies that the candidates who win against women and against men are similar and the difference in candidate ideology now only arises from facing the opponent whose gender is as good as randomly assigned as a result of the RDD setting. This effectively shuts down the second potential channel of impact and allows us to simply compare the winners who barely faced female opponents and who barely faced male opponents using equation (3), and estimate the impact of opponent gender on candidate ideology.

For the second part of the analysis, I relax the assumption that the opponent gender has no impact on the candidate probability of winning. The goal then becomes to estimate the
impact of opponent gender on candidate ideology conditional on winning the race, and I do so by making assumptions about the unobserved political ideology of the losers and providing bounds on the potential treatment effect. In the remainder of this section, I discuss the identifying assumption of my model and explain the empirical design of the two analyses mentioned above.

1. Identification assumption

The underlying assumption of the regression discontinuity design embedded in equation (1) is that all candidate characteristics change continuously at the threshold while there is a discrete change in the opponent’s gender. If the candidates can sort across the discontinuity at the threshold, this implies self selection into treatment and will threaten the validity of the assumption. A canonical check for RD design is to use the McCrary test (McCrary 2008) where the key idea behind the test is that in the absence of manipulation around the cutoff, the density of the units should be continuous around the cutoff point. I run a balance check on series of pre-treatment variables of the candidates and see if there’s a sudden jump at the threshold.

For variables including candidate gender, party, victory margin at the primary, number of candidates, candidate vote share, and the vote share of the runner up at the primary, I regress the treatment variable on these variables and use the coefficients from this regressions to predict the treatment status for each candidate. Then I test whether the predicted treatment status jumps at the threshold. To avoid dropping observations, I include a dummy equal to one when the variables vote percentage at the primary and the victory margin at the primary are missing and replace the missing values with zeros. The result is shown in Figure VI, where I fit a quadratic polynomial on each side of the threshold. Figure VI shows the lack of any jump at the cutoff point for the predicted treatment status. The coefficient for the local linear RD point estimator is 0.004 with a standard error of 0.017. I also individually examine whether there is a discontinuity in any of the above variables used to predict treatment status. The corresponding graphs and tables are included in Appendix A. Overall, none of the six coefficients is statistically significant. With these general balance test results, I cautiously argue that that there is no systematic sorting of candidates at the
threshold.

The balance test may have shown us that the candidates at the threshold may be similar. But the fact that the candidates facing women versus men are similar at the threshold does not imply the candidates who win against women versus men are similar as well. As my outcomes of interest – ideology scores calculated using candidates’ voting pattern later in Congress – is measured only after the candidates have been elected, if an effect exists, I cannot distinguish if the effect is driven by running against women or by changing the pool of winning candidates who ran against women. The following next two subsections discuss the estimation strategies to isolate the effect of running against women.

2. Estimation without bounds

*Impact on winning* – I run the following regression to study the impact of opponent gender on the candidate’s probability of winning:

\[
W_{it} = \beta_0 + \beta_1 D_{it} + \beta_2 X_{it} + \beta_3 D_{it} X_{it} + \epsilon_{it}
\]  

(4)

where \(W_{it}\) is a binary outcome of candidate \(i\) in election \(t\) which equals 1 if \(i\) won the race and 0 otherwise, \(D_{it}\) is a binary indicator that equals 1 if \(i\)’s opponent in election \(t\) is a female. \(X_{it}\) is the running variable that measures the difference between the vote share of the female candidate and the male candidate in \(i\)’s opponent’s primary election such that if \(X > 0\), then \(D = 1\) and if \(X < 0\), then \(D = 0\). Table IV and Figure VII presents the impact of running against women on the candidate’s likelihood of winning the race.

Leaving columns 2 and 3 for discussions later in the paper, we see from column 1 in Table IV that running against women has a minimal impact on the probability of winning the race. Compared to when running against a male opponent, facing a female opponent decreases your chances of winning by 0.3 pp on average (not significant). This allows for a simple comparison between the winners who ran against close female opponent and those won ran against close male opponent, using equation (3).
3. Bounding strategy

Even though the gender of the opponent seems to have near zero and statistically insignificant effect on the candidate probability of winning the race, studying the impact separately for the two parties in columns 2 and 3 of Table IV tells a different story. Indeed, the near zero impact is masking a heterogeneous impact on winning by the two parties; Though insignificant, Republican candidates are 16 pp less likely to win against female opponents versus male opponents and Democratic candidates are 11 pp more likely to win against female opponents versus male opponents.

I no longer assume that opponent gender has no impact on the candidate probability of winning and acknowledge that the impact of facing female opponent on legislator ideology can result both from the impact on the ideology conditional on winning and the change in the probability of winning. Then the goal of this section is to estimate the impact on candidate ideology conditional on winning the race, which will shut down whatever effect that comes from the fact that perhaps certain candidates are more likely to win against women. As will be shown later in the paper, the results from the non-bounding method in section 2. show that the impact exists only for the Democratic party and I focus my analysis for this section exclusively on the Democratic party.

In order to disentangle the two potential channels and tease out the impact from candidate facing female opponent conditional on winning, I use a bounding method with the regression discontinuity design. Bounding method with RDD is useful when between the timing of treatment and of the outcome, there exists another ‘treatment’ that people self-select into that may potentially affect the population for which the outcome is available, thereby making the comparison of only those with the non-missing outcomes biased.

For example, Anagol and Fujiwara (2016) study whether runner-ups have higher chances of winning the race than third-ranking candidates. Because close second-place candidates and close third-place candidate still have to decide whether to run in the final round, the candidates who end up running the final round, and finally win, may be different from those who dropped out before the final round. The decision to run in the final round makes the sample with the outcome selective, and a simple comparison of close second-place candidates’
and close third-place candidates’ outcome is now a blend of both the impact of their ranks on the outcome, and their decision to run in the final round. Using the bounds with RDD, the authors show that close second-place candidates are substantially more likely than close third-place candidates to win the race, conditional on staying in the race.

Similarly, Granzier, Pons, and Tricaud (2021) also use the bounds method with RDD to show the effect of candidate rankings on the probability of winning, and show that candidate ranking in the first round of race affects the probability of winning in the final round for not only the 1st vs 2nd ranking candidates, but also 2nd vs 3rd, and 3rd vs 4th candidates, conditional on staying in the race.

Studying the impact of women running the race on the candidates’ political orientation later in Congress poses a similar challenge. Between running in the race and voting in Congress, the candidates must win the race. And if candidates who tend to win against women are innately different from candidates who tend to lose against women, this will show up as differences in the political orientation in Congress, which I could falsely ascribe to a change caused by the female opponents they face. Bounding strategy with RDD helps to isolate the impact on candidate ideology that comes from running against women, conditional on the candidate winning the race. The rest of the section builds on Lee (2009) and follows closely the framework of Anagol and Fujiwara (2016) and of Granzier, Pons, and Tricaud (2021).

Using the potential outcomes framework, I first define $W_0$ and $W_1$ as indicator variables of winning the general election when $D = 0$ (candidate faces men) and $D = 1$ (candidate faces women), respectively. In the data, I only observe $W = DW_1 + (1 - D)W_0$: I know whether the candidate who faced woman wins the general election but not whether the same candidate would have won if s/he faced man, and vice versa. Next, I define $Y_0$ and $Y_1$ as the political ideology scores of the legislators, conditional on winning when $D = 0$ and $D = 1$, respectively. I only observe $Y = WDY_1 + (1 - D)Y_0$. When the candidate doesn’t win ($W = 0$), s/he does not vote in Congress and I do not observe what their ideology would have been if s/he had won. When the candidate does win ($W = 1$), I observe the ideological scores of the candidate who faced woman but not what her/his scores would have been if had faced man, and vice versa.
Next, I define four types of candidates: “always takers”, who always wins in the general election, whether they face woman or man; “never takers”, who never win the general election; “compliers”, who win if faced woman but not man; and “defiers”, who win if faced man but not woman. My key assumption, which follows Lee (2009), is that there are no defiers; all candidates who faced men and won the election would also have won the election if they faced women. The variables $D$, $W_0$, $W_1$, $Y_0$, and $Y_1$ can be thought of as functions of the candidate and the running variable $x$ and their limits at the cutoff ($x = 0$) can be estimated.

We want to estimate the impact on the ideology conditional on winning the election

$$E[Y_1 - Y_0|x = 0, W_1 = 1] = E[Y_1|x = 0, W_1 = 1] - E[Y_0|x = 0, W_1 = 1] \quad (5)$$

Omitting the $x = 0$ condition hereafter, the first term on the right side can be expressed as:

$$E[Y_1|W_1 = 1] = \frac{E[Y_1|W_1 = 1]Pr(W_1 = 1)}{Pr(W_1 = 1)} = \frac{E[Y_1W_1]}{Pr(W_1 = 1)} \quad (6)$$

The second term can be expressed similarly as:

$$E[Y_0|W_1 = 1] = \frac{E[Y_0|W_1 = 1]Pr(W_1 = 1)}{Pr(W_1 = 1)} = \frac{E[Y_0W_1]}{Pr(W_1 = 1)} + \frac{E[Y_0W_1 > W_0]Pr(W_1 > W_0)}{E[W_1]} \quad (7)$$
Putting equation (3) and (4) together, we get:

$$
E[Y_1 - Y_0|W_1 = 1] = \frac{1}{E[W_1]}(E[Y_1W_1 - Y_0W_0] - E[Y_0|W_1 > W_0]Pr(W_1 > W_0))
$$

(8)

Placing back the $x = 0$ condition, the effect on the conditional outcome can be written as:

$$
\frac{1}{E[W_1|x = 0]} \left( E[Y_1W_1 - Y_0W_0|x = 0] - \lim_{x \downarrow 0} E[W|x] \cdot E[Y_0|x = 0, W_1 > W_0] \right)
$$

(9)

Every term on the right hand side can be estimated, except for the unobservable term. The first term on the right hand side $E[W_1|x = 0]$, which is the probability that a female facing candidate would win the race, and the third term $\text{Prob}(W_1 > W_0|x = 0)$, the impact on winning the race, are both estimated in Subsection 2. The second term, $E[Y_1W_1 - Y_0W_0|x = 0]$, is the impact on the unconditional outcome and is estimated using equation (1). This term is labeled “unconditional” because the outcome is defined for everyone in the sample, regardless of whether they won the race or not. In theory, this term measures the impact of opponent’s gender on all candidate’s ideology, regardless of whether she or he won the race.

I use the DW-Nominate scores for the outcome of the winners. Then I arbitrarily choose a number to assign as an outcome for all losers of the race, because the losers do not get to vote in Congress and thus we do not observe their outcome, $Y_{it}$.

This arbitrary choice seems counterintuitive at a first pass. What does it mean to choose and assign a random number as the outcome of all the losers when these numbers have specific meanings? For example, if I arbitrarily assign 1 as the outcome of all the losers, it appears as if I’m forcibly making all losers to be extremely conservative. If I instead arbitrarily assign $-1$ to all the losers, it’s as if I’m making all the losers to be extremely liberal. What if I plug in $-2$? 10? Indeed, if I were to plug in different arbitrary numbers and run equation (1), this will produce different estimates for the effect on the unconditional outcome. But let us remember that the goal of this paper is not to recover the impact on the unconditional outcome, but rather the impact on the conditional outcome. A closer look at
equation (9) shows us that the unconditional outcome is a part of the conditional outcome. This renders the choice of the arbitrary value innocuous in calculating the impact on the conditional outcome, because as I change the arbitrary value the impact on the unconditional outcome also changes, eventually offsetting each other when both changes are reflected in estimating (9). This is shown more in detail in Appendix B.

Finally, the last term of equation (9) is the average outcome of the close compliers if they had won the election, absent treatment (i.e., when they faced male opponent). By definition, compliers do not win when they run against men but only when they run against women and thus this term is unobservable. But since all other terms are observable, we can derive the bounds on the treatment effect conditional on winning by making assumptions about this unobservable term. To obtain an upper bound for the conditional outcome, I set $E[Y_0|x = 0, W_1 > W_0] = -1$, because the largest possible effect of running against a female opponent on the ideological score conditional on winning the election, happens if we assume that barely male-facing compliers are extremely liberal. I also use a more conservative upper bound by replacing the unobservable term by the 10th and 25th percentile political scores of the winners, allowing the barely male-facing compliers to be a little more conservative.

To obtain a lower bound, I replace the unobservable term by the average political scores of barely female facing candidates who win the election. I also use a more conservative lower bound by replacing the unobservable term by the top 75th and 90th percentile political scores of the winners, allowing the barely male-facing compliers to be more conservative.

Finally, I use a bootstrapping procedure to estimate the standard errors of the bounds: I draw a sample from the districts with replacement, compute the lower and upper bounds as indicated above, repeat the two steps 1,000 times, and estimate the empirical standard deviation of both bounds. Next section presents the results.

V. Results and Discussion

This section presents the empirical results of the two estimation methods, first without and later with the bounds. Under the assumption that the gender of the opponent does not affect candidate probability of winning the race, Tables V and VI show the result of
running regression (3) with outcomes DW-Nominate scores in the first and second dimension, respectively. We see from Table V that running against female opponent has no impact on the first dimensional scores that map the liberal conservative scale. Both parties show insignificant and near zero effect on the ideology when they run against female opponent versus male. This is consistent with the short term finding where we find null effect using the CF scores to study the impact of opponent gender on candidate ideology during the electoral campaign, where the CF scores map onto the liberal-conservative (first) dimension of Nominate scores with a bivariate correlation of $r = 0.92$ (Bonica 2013).

But while I find no impact on the first dimensional scores, I find statistically significant impact on the second dimension of the DW-Nominate scores. Table VI column 1 shows that facing a female opponent increases legislator Nominate scores in the second dimension by 0.17 on average (significant at 5% level), meaning that legislators who ran against female voted more conservatively on racial and cultural issues later in Congress, compared to legislators who ran against male. Looking at the effects separately by two parties in columns 2 and 3, we see that this effect is mostly driven by legislators in the Democratic party, whose score increases by 0.27 (significant at 1% level). With a mean score 1.09 and standard deviation of 0.38 for male-facing Democratic legislators in the sample, the 0.27 treatment effect translates to a 0.71 standard deviation increase in the ideology score.

We now relax the assumption that the gender of the opponent does not affect candidate probability of winning the race. Table VII provides the resulting bounds and bootstrapped standard errors of the effects of facing female opponents on the legislators’ Nominate scores in the second dimension for the Democratic party. Under the assumption that male-facing complier candidates would have assumed the average orientation of candidates who won against female contender, facing a female candidate increases a legislator’s political orientation by 0.218 (significant at 5% level). Using a more conservative assumption that the male-facing compliers would have had scores equivalent to the 75th percentile of the Democratic Congressmen, this effect drops to 0.213 (significant at 5% level), and with even more conservative assumption (now 90th percentile) drops to 0.180 without any statistical significance. As chances of all male-facing compliers having a 90th percentile score of elected Democrat officials is highly unlikely, I now focus on interpreting the results using the con-
servative bounds. How does one interpret the effect size of 0.214-0.307? To get an idea of the magnitude of this change, let’s look at the overall standard deviation of the Nominate score in the 2nd dimension within the Democratic party Congressmen who run against male: 0.438. This gives us the spread of the score within the party across all Democratic Congressman who ran against male opponent in 1956–2014. Dividing the effect size 0.214 and 0.307 by 0.422, we get a 0.50 to 0.73 standard deviation increase in the political score for candidates who ran against women.

V.A. Second dimension DW-Nominate scores

Let us now delve deeper into what the second dimension Nominate scores measure. Though the two dimensions of the Nominate scores have collapsed into unidimensional liberal-conservative space in modern days, when important, the second dimensional ideology score captures the candidate ideology on the racial and cultural issues (Poole and Rosenthal 2001). And while the first dimension is supposed to give us the left right divide between the two parties, the second dimension picks up regional differences within the parties, such as between Northern and Southern democrats.

Figure VIII shows a graph that plots every member of House in each congress from 1956 to 2014. The two linear regression lines show us the correlation between the first dimension score and the second within each party. The higher first dimension scores are correlated with higher second dimension scores within the party, meaning that more conservative politicians within each party tends to have higher second dimension scores. Figure IX is excerpted from Poole and Rosenthal (2001), where they plot the time trend of the mean second dimension scores of each party. On the Y axis of the graph is the “North” and “South” divide, and we see that the Southern Democrats have higher scores while the Republicans have lower scores in this second dimension.

What, then, is a “Southern Democratic” way of thinking? The Southern Democrat is a historic faction of Democratic party and were strongly pro-slavery, pro-states’ rights, and pro-free trade. During 1960s and 70s, the southern Democrats often voted Republican, driven not only by the issue of race, but also by white evangelical Christians’ opposition to abortion and other “culture war” issues. Slowly, white working-class and rural bases of the
Southern Democrats switched to Republican and by 2010 Southern Democrats ceased to be a major faction of the Democratic party.

The general consensus is that the two-dimensional dynamic spatial model is the best fitting model for Congresses 1–99 and that the second dimension slowly evaporated since the 2000s. As late as 1990s, the second dimension picked up differences within each parties over abortion, gun rights, and other social or lifestyle issues where the roll call votes on these issues often cut through the parties along the second dimension (Poole 2017) and in recent years, the second dimension serves to pick up votes that do not fit the typical liberal-conservative dimension. A good example of a one-dimensional fit vote is 2010/2011 votes on the Patient Protection and Affordable Care act and for a two-dimensional fit vote would be the 1964 vote on the Civil Rights Act, before race was absorbed into the first dimension (Hare and Poole 2014). Figure XI reverberates this claim in which thereafter 2010 we see the crossing of the two parties in this dimension. I rerun the main analysis in Appendix C as a robustness check to show that the results don’t change even if I restrict the sample to time periods before 2010.

V.B. Conservative voting in social dimension

With this understanding of the second dimension score, what does the results in Table VII tell us? First, that a personal identity – gender – of the opponent leaves a lasting impact on the legislative voting behavior of the candidate in the long term. The gender of the opponent may not impact the candidate (legislator) in the popular liberal-conservative scale, but the impact exists on a racial and cultural margin and is significant. The implication of this finding is encouraging; it tells us that by simply participating in the race, one has a potential to shape the dialogue of policy and lawmaking in Congress even when they do not win the race.

Candidates’ response and adopting of their behavior to the identity of the opponent implies that the data tells a story more aligned with the spatial model of Downs (1957), Calvert (1985), and Wittman (1977) than with the citizen-candidate model of Besley and Coate (1997), Osborne and Slivinski (1996), and Besley and Case (2003), where the sole strategic action in the citizen-candidate model is captured by the candidates’ decision to enter the race. Without a full policy commitment as in the Downsian model, the citizen-
candidates essentially offer their ideal points to the voters by entering the race and carry out their preferred policy, regardless of who the opponents are, after being appointed to the office. On the other hand, Downsian spatial models show candidates adapting their issue space as a response to the opponent in order to win the race.

Second, that running against female candidates make Democratic legislators vote more culturally and racially conservatively while in office. As can be seen in Figure XI, this means that the Democratic legislators pull away from the center and adopt more culturally extremist stances when running against Republican women, who are located further away from the center.

What can we say about the direction of the impact? I list a few potential but non-comprehensive reasons below. The first natural reason may be that the candidates campaign differently against a female opponent and are now simply following up on the promises they’ve made during the election. Studying how the candidates’ emphasis on policy issues and their campaign strategies shape the subsequent legislative behavior is reasonable, as the issues candidates discuss during their campaigns predominantly end up as the issues they legislate on (Mayhew 1974; Sulkin 2009, 2011). The challenge of supporting this claim with causal evidence lies on the need for data on electoral campaigns such as speech, ads, and/or campaign websites of the candidates and more importantly ability to parse through them to isolate the issues that fall into the second dimension of the Nominate score. While such an endeavor is beyond the scope of this paper, I look at past literature on candidate interactions and issue positioning as part of electoral campaigns to shed light on the relationship between opponent gender and electoral politics.

The unifying conclusion of various theories on campaigning including saliency theory (Budge and Farlie 1983), emphasis allocation theory (Page 1976), intentional priming (Jacobs and Shapiro 1994), and issue ownership (Petrocik 1996) is that the candidates attempt to emphasize the issues that work to their advantage and deemphasize those that do not.\footnote{For example, Republicans are perceived to be better equipped to handle issues related to foreign policy/national defense and economy while Democrats are perceived to be stronger with social welfare issues such as healthcare and education. Candidates from both parties would try to steer the conversation and the focus of the election to the issues their own party has expertise in.} But this may not necessarily be the case for how the candidates strategize their electoral
campaigns against different gendered opponent. Although female candidates are assumed to have policy advantage on traditionally “feminine” issues, the female candidates themselves would prefer to run as the ‘candidate for office’ instead of ‘female candidate for office’ (Larson 2001; Witt, Paget, and Matthews 1994). They therefore may strategically choose to stress “masculine” issues during their campaign so as to convince the voters of their all-around policy expertise and competence. Similarly, male candidates who run against female opponents have a choice of diverging from women’s issues and focusing on what the public would typically perceive as more masculine, or strategically campaigning on feminine issues to 1) appeal to women voters and 2) to make the female opponents respond and consequently prime gender stereotypes to the electorate (Fox 1997, 2000; Windett 2014).

While a plethora of papers study gendered differences in candidates’ political campaign and discussion of policy issues, a few handful of papers offer insights on how the candidates respond to opposite gendered opponents during the campaign. Dolan (2008) studies Congressional candidate campaign websites of male candidates who ran against women in years 2002 and 2006 to find no evidence that men “play” to women voters through their attention to women’s issues while Meyer (2018) uses the social media of 2014 and 2016 Congressional candidates who faced opposite gendered opponents in the general election to find that the male candidates decrease their attention to women’s issues when faced with a male opponent versus a female opponent. Similarly, Fox (1997, 2000) uses in-depth interviews and Windett (2014) TV ads to find that male candidates running against female opponents strategically stress feminine issues at the onset of the campaign. Though they all suffer from a small sample size and from self selection, the papers nonetheless bring valuable insight into how the candidates in Congressional races in 2000s correspond with opposite gendered opponent on their political platforms.

Second, the candidate may be able to move closer to his ideal point without fear of losing the support base if the opponent is also further enough from the center, and if the candidate believes he possesses an upper hand compared to the opponent. Valence theorem, which incorporates the idea into the Downsian spatial model that on top of competing on policy issues, candidates also compete on nonpolicy characteristics that are universally desirable such as incumbency, competence, integrity, higher intelligence or charisma etc, is well suited
to explain this phenomena. The valence theorem predicts that candidates will strategically take more extreme positions when they are disadvantaged on a separate dimension other than policy. Ansolabehere and Snyder (2000) and Groseclose (2001) argue that if the candidates have valence advantage, this affect the way politicians position themselves in a policy space. The model predicts high valence candidates to take moderate positions and for low valence candidates to take more extreme positions. This is because by moving closer to the center, the candidate with higher valence can deemphasize the importance of policy and consequently increase the importance of his valence advantage to the voters.

But a more general consensus of the Downsian spatial models that incorporate valence is that if a candidate has a large enough valence advantage, then s/he has considerable freedom in policy position while still leaving the opponent in no winning location (Ansolabehere and Snyder 2000). That is, when a valence disadvantaged opponent pulls away from the center, the candidate can move closer to his ideal point and still maintain the support base. Gender, as a universally observable characteristics of candidates, is a dimension of valence the candidates can compete on. The theorem would predict that if a candidate believes his gender acts as an advantage – or his opponent’s gender acts as a disadvantage to the opponent – he will move closer to the center to converge on the policy issues and to steer the electorate’s attention to his gender. On the other hand, if this advantage is large enough, it would grant the candidate freedom to vote closer to his ideal point instead and may instead pull him further away from the centrist position.

Figure XI shows that Republican females are located at a more extreme position relative to Republican males in their DW-Nominate scores in the second dimension. Coupled with the fact that Republican women candidates perform the worst during the time period of study, Democratic legislators may be able to afford to move away from the center and closer to their ideal point when running against Republican female opponent and still maintain the support of the voter base as a result. This may also explain why we do not see any effect for the Republican legislators running against Democratic female opponents. Democratic female legislators as a matter of fact stand closer to Republican male legislators compared to their male copartisans, and with Democratic females performing as well as – sometimes better – than their male copartisans, the Republican candidates can not afford to move away
from the center when s/he faces a female opponent versus a male opponent.

Finally, the Democratic legislators could be pulling away from the center as result of uncertainty. When a Republican female enters the race, her gender acts much more of a wild card than facing a Republican male opponent, and the Democratic candidate might as well choose his/her ideal point, since nothing is predictable and converging doesn’t necessarily help winning the race.

VI. Conclusion

While the notion that an opponent influences candidate’s political behavior appears to be obvious, to this point and to my knowledge, there has been no empirical evidence to causally identify and quantitatively substantiate this intuition. This paper begins to fill the gap in the literature. I first show that while the competition itself makes the candidate diverge away from the center and closer to party extreme, the gender of the opponent does not necessarily shift candidate ideology on a liberal-conservative scale before the election.

Instead, I find that the opponents leave a lasting impact on future Congress by changing the voting pattern of the candidates on racial and cultural issues. Using a regression discontinuity design combined with a bounding strategy in the congressional election setting, I find that Democratic legislators who run against females later score 0.21 higher in their DW-Nominate scores in the second dimension, compared to legislators who do not, meaning that female-facing Democratic candidates show more conservative stance when voting for racial, cultural, and other social issues in Congress, conditional on having won the race.

This is an important and novel result. It shows that even if one loses the race, the act of participating in the race itself affect other candidates’ political orientation and voting pattern later in Congress. This may be due to campaign promises made before the election while facing the opponent or because the legislators don’t find the opponent to be a threat enough, allowing themselves to position closer to their ideal point. The implication of this work to the political world is meaningful; when one decides to run in a race, her decision to run in itself may leave a lasting impact on future policy even if she loses and it is likely her personal traits and preferences that drive the change.
We know that when female candidates run, they can win, and when in office they better represent the interest of same gendered voters and focus on different platforms compared to their male colleagues as documented by previous studies. But encouraging and facilitating more entry of female candidates into the political arena has an impact that goes beyond them winning the elections. They influence the platforms and the ex post behavior of the winner of the election by simply participating in the dialogue and thereby shaping legislative voting pattern of the winner. Though this paper leaves us with a sobering finding with an unintended consequences of a well-intended push for political representation, understanding that the effort at representation itself can bring a significant change is consequential.
FIGURE I. Fraction of female candidates in U.S. House of Representative general election

Notes: Sample is candidates from two major parties who ran in the general election of the US house of representatives race in 1956–2014.
FIGURE II. CF scores time trend

Notes: Above figure shows the time trend of the campaign finance scores of general election candidates running for the U.S. House of Representatives in 1980–2014.
FIGURE III. CF scores histogram

Notes: Above figures plots the histogram of campaign finance scores of the general election candidates running for the US house of representatives in 1980–2014. The higher scores correspond to more conservative ideology.
FIGURE IV. Time trend of Nominate scores

Notes: Above figures show the time trend of Nominate scores in the first (left) and second (right) dimension of the US house of representatives in office 1956 – 2014. The scores in the first dimension quantifies the typical left-right (liberal-conservative) spectrum on the economic matters while scores in the second dimension represent attitudes on salient cultural and lifestyle issues of the day. Scores closer to 1 are described as conservative whereas scores closer to –1 can be described as liberal. Score close to zero is described as moderate.
FIGURE V. Histogram of Nominate scores

Notes: Above figures show the histogram of Nominate scores in the first(left) and second(right) dimension of the US house of representatives in office 1956 - 2014. The scores in the first dimension quantifies the typical left-right (liberal-conservative) spectrum on the economic matters while scores in the second dimension represent attitudes on salient cultural and lifestyle issues of the day. Scores closer to 1 are described as conservative whereas scores closer to -1 can be described as liberal. Score close to zero is described as moderate.
FIGURE VI. General balance test

Note: Sample is U.S. House general election candidates from two major parties in 1956–2014. I further restrict the sample to elections where the opponent’s primary’s top two contestants are of opposite gender. Running variable is the margin of victory of female candidate to male candidate in the opposing party’s primary election. Y-axis is the predicted treatment status of the candidate where the treatment is facing a female opponent in the general election. Each dot indicates the average value of the outcome within certain bins of the running variable. Observations corresponding to female-facing candidates are on the right side of the threshold while those corresponding to male-facing candidates are on the left side of the threshold.
FIGURE VII. Impact on winning

Notes: Sample is U.S. House general election candidates from two major parties in 1956–2014. I further restrict the sample to elections where the opponent’s primary’s top two contestants are of opposite gender. X-axis is the difference of female candidate’s vote share to that of male candidate in the opponent’s primary election and y-axis is the probability of winning the race when running against a female opponent – the case of candidates placed in the right half of the graph – versus when running against a male opponent – the left half of the graph.
FIGURE VIII. Scatter plot of Nominate scores

Notes: The graph plots every members of House serving in Congress from 1956 to 2014. Each point represents a legislator’s Nominate score and the figure may contain multiple points for a unique legislator if s/he served multiple terms in Congress. X-axis is the DW-Nominate scores in the first dimension and y-axis is the DW-Nominate scores in the second dimension. The two linear regression lines depict the correlation between the first dimension score and the second within each party.
FIGURE IX. Party means of 2nd dimension Nominate Scores
Notes: The graph is sourced from Poole and Rosenthal (2001) and plots the historical party means of the Nominate scores in the second dimension.
FIGURE X. Nominate 1st score time trend
Notes: The graph plots means of the first dimension Nominate scores of Congressmen and women serving in Congress from 1956 to 2014 by party and by gender.
FIGURE XI. Nominate 2nd score time trend
Notes: The graph plots means of the second dimension Nominate scores of Congressmen and women serving in Congress from 1956 to 2014 by party and by gender.
### Table I. Summary statistics of DW-Nominate scores

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs</td>
<td>Mean</td>
<td>SD</td>
<td>Obs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1st dimension</td>
<td>12,431</td>
<td>-0.038</td>
<td>0.388</td>
<td>5,531</td>
<td>0.354</td>
<td>0.183</td>
</tr>
<tr>
<td>2nd dimension</td>
<td>12,431</td>
<td>0.018</td>
<td>0.419</td>
<td>5,531</td>
<td>-0.129</td>
<td>0.341</td>
</tr>
</tbody>
</table>

Notes: Sample is U.S. House of Representatives in 1956-2014.
TABLE II. Short-term impact of female opponent on candidate ideology

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Republican</th>
<th>Democratic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female opponent</td>
<td>0.058</td>
<td>0.032</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td>(0.158)</td>
<td>(0.094)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,090</td>
<td>618</td>
<td>472</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.262</td>
<td>0.281</td>
<td>0.240</td>
</tr>
<tr>
<td>Observations left</td>
<td>273</td>
<td>147</td>
<td>126</td>
</tr>
<tr>
<td>Observations right</td>
<td>247</td>
<td>143</td>
<td>105</td>
</tr>
</tbody>
</table>

Note: The unit of observation is candidate × year. Sample includes both the winners and the losers of the U.S. House of Representatives general election in 1980 – 2014. The sample is further restricted to candidates whose opposing party’s primary election’s top two candidates were of different gender. The treatment variable ‘Female opponent’ equals 1 if the candidate faced a woman in the general election and 0 if man. Outcome variable is candidates’ ideology measured by the campaign finance (CF) score, where the higher CF scores corresponds to more conservative stance on the liberal-conservative scale. I use triangular kernel that weights points near the threshold more than those far from the threshold and use the preferred choice of polynomial order of 1, which gives the standard local linear RD point estimator. The bandwidth are computed following the procedure by Calonico et al. (2017), which trade-off bias and variance of the treatment effect to find the MSE-optimal bandwidth. Standard errors are clustered at the district level and are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
TABLE III. Suggestive evidence on the effect of competition on candidate ideology

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Both parties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contested</td>
<td>0.005</td>
<td>0.006</td>
<td>-0.021***</td>
<td>-0.015**</td>
<td>0.032***</td>
<td>0.024***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>0.001</td>
<td>0.097***</td>
<td>0.097***</td>
<td>-0.081***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.008)</td>
<td></td>
<td></td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8,827</td>
<td>8,827</td>
<td>4,426</td>
<td>4,426</td>
<td>4,147</td>
<td>4,147</td>
</tr>
</tbody>
</table>

**Panel A: Opponent exists**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate voteshare</td>
<td>0.052</td>
<td>0.055</td>
<td>0.569***</td>
<td>0.434***</td>
<td>-0.501***</td>
<td>-0.342***</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.037)</td>
<td>(0.051)</td>
<td>(0.051)</td>
<td>(0.055)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>-0.002</td>
<td>0.076***</td>
<td>0.076***</td>
<td>-0.062***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
<td></td>
<td></td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>7,704</td>
<td>7,704</td>
<td>3,760</td>
<td>3,760</td>
<td>3,655</td>
<td>3,655</td>
</tr>
</tbody>
</table>

**Panel B: Candidate voteshare**

Note: The unit of observation is candidate × race (year) and the outcome variable is candidates’ ideology measured by the campaign finance (CF) score, where the higher CF scores correspond to more conservative stance on the liberal-conservative scale. Sample is candidates who ran in the U.S. House general elections twice or more between 1980–2014. Variable Contested equals 1 if the candidate has an opponent in the race at year $t$ and variable Incumbent equals 1 if the candidate is an incumbent in the race at year $t$. Candidate voteshare is the percentage of the votes the candidate received in the race $t$. All models include candidate FE and state-year FE. Robust standard errors clustered at the candidate level are in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opponent is female</td>
<td>-0.034</td>
<td>-0.160</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>(0.084)</td>
<td>(0.109)</td>
<td>(0.116)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,473</td>
<td>864</td>
<td>609</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.184</td>
<td>0.180</td>
<td>0.253</td>
</tr>
<tr>
<td>Observations left</td>
<td>264</td>
<td>133</td>
<td>164</td>
</tr>
<tr>
<td>Observations right</td>
<td>237</td>
<td>135</td>
<td>129</td>
</tr>
</tbody>
</table>

Note: The unit of observation is candidate × year. Sample includes both the winners and the losers of the U.S. House of Representatives general election in 1956 – 2014. The sample is further restricted to candidates whose opponent’s primary election’s top two candidates were of different gender. Outcome variable is a binary that equals 1 if the candidate won the race. I use triangular kernel that weights points near the threshold more than those far from the threshold and use the preferred choice of polynomial order of 1, which gives the standard local linear RD point estimator. The bandwidth are computed following the procedure by Calonico et al. (2017), which trade-off bias and variance of the treatment effect to find the MSE-optimal bandwidth. Standard errors are clustered at the district level and are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
TABLE V. Impact on DW-Nominate score in 1st dimension without bounds

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opponent is female</td>
<td>-0.003</td>
<td>0.010</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.058)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>Observations</td>
<td>833</td>
<td>468</td>
<td>365</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.275</td>
<td>0.220</td>
<td>0.286</td>
</tr>
<tr>
<td>Observations left</td>
<td>251</td>
<td>111</td>
<td>125</td>
</tr>
<tr>
<td>Observations right</td>
<td>235</td>
<td>119</td>
<td>103</td>
</tr>
</tbody>
</table>

Note: The unit of observation is candidate by year. Sample is the winners of US house of representatives election from the two major parties in 1956 - 2014 whose opponent’s primary’s top two candidates were of different gender. I use triangular kernel that weights points near the threshold more than those far from the threshold and use the preferred choice of polynomial order of 1, which gives the standard local linear RD point estimator. The bandwidth are computed following the procedure by Calonico et al. (2017), which trade-off bias and variance of the treatment effect to find the MSE-optimal bandwidth. Standard errors are clustered at the district level and are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
### TABLE VI. Impact on DW-Nominate score in 2nd dimension without bounds

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Republican</td>
<td>Democratic</td>
</tr>
<tr>
<td>Opponent is female</td>
<td>0.166**</td>
<td>0.073</td>
<td>0.272***</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.099)</td>
<td>(0.097)</td>
</tr>
<tr>
<td>Observations</td>
<td>833</td>
<td>468</td>
<td>365</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.217</td>
<td>0.228</td>
<td>0.182</td>
</tr>
<tr>
<td>Observations left</td>
<td>212</td>
<td>116</td>
<td>87</td>
</tr>
<tr>
<td>Observations right</td>
<td>200</td>
<td>121</td>
<td>77</td>
</tr>
</tbody>
</table>

Note: The unit of observation is candidate by year. Sample is the winners of US house of representatives election from the two major parties in 1956 - 2014 whose opponent’s primary’s top two candidates were of different gender. I use triangular kernel that weights points near the threshold more than those far from the threshold and use the preferred choice of polynomial order of 1, which gives the standard local linear RD point estimator. The bandwidth are computed following the procedure by Calonico et al. (2017), which trade-off bias and variance of the treatment effect to find the MSE-optimal bandwidth. Standard errors are clustered at the district level and are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$
TABLE VII. Bounds on the impact on political ideology for Democratic party, conditional on winning

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regular bound</td>
<td>Conservative</td>
<td>More conservative</td>
</tr>
<tr>
<td>Lower bound</td>
<td>0.219**</td>
<td>0.214**</td>
<td>0.182</td>
</tr>
<tr>
<td>Boot. std error</td>
<td>(0.100)</td>
<td>(0.108)</td>
<td>(0.130)</td>
</tr>
<tr>
<td>Upper bound</td>
<td>0.385**</td>
<td>0.307***</td>
<td>0.277***</td>
</tr>
<tr>
<td>Boot. std error</td>
<td>(0.166)</td>
<td>(0.105)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.930</td>
<td>0.930</td>
<td>0.930</td>
</tr>
</tbody>
</table>

Notes: The outcome variable is the Nominate score in the 2nd dimension for the Democratic party candidates. Regular upper bound assumes that the barely male-facing compliers are most conservative and regular lower bound assumes them to have the average political scores of barely female-facing winners. Conservative and more conservative upper bounds assume that barely male-facing compliers have the 10th and 25th percentile Nominate scores of the winners, respectively, while the conservative and more conservative lower bounds assume that they have the 75th and 90th percentile Nominate scores of the winners, respectively. The mean is the average value of the ideology score for the male-facing candidates at the threshold, conditional on winning the race. I use triangular kernel that weights points near the threshold more than those far from the threshold and use the preferred choice of polynomial order of 1, which gives the standard local linear RD point estimator. The bandwidth are computed following the procedure by (Calonico et al. 2017), which trade-off bias and variance of the treatment effect to find the MSE-optimal bandwidth. The bootstrapped standard errors are clustered at the district level and are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
References


### Appendix A

#### Table A.1. Balance test

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome variables</strong></td>
<td>Female Party</td>
<td>Victory margin in primary</td>
<td>Candidate number in primary</td>
<td>Voteshare in primary</td>
<td>Runner-up voteshare in primary</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td>0.018</td>
<td>-0.044</td>
<td>-0.052</td>
<td>-0.338</td>
<td>0.069</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.089)</td>
<td>(0.043)</td>
<td>(0.273)</td>
<td>(0.045)</td>
<td>(0.025)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>1,473</td>
<td>1,473</td>
<td>1,473</td>
<td>1,473</td>
<td>1473</td>
<td>774</td>
</tr>
<tr>
<td><strong>Polynomial order</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>0.310</td>
<td>0.230</td>
<td>0.245</td>
<td>0.264</td>
<td>0.215</td>
<td>0.238</td>
</tr>
<tr>
<td><strong>Observations left</strong></td>
<td>395</td>
<td>329</td>
<td>342</td>
<td>358</td>
<td>306</td>
<td>175</td>
</tr>
<tr>
<td><strong>Observations right</strong></td>
<td>361</td>
<td>286</td>
<td>299</td>
<td>319</td>
<td>270</td>
<td>151</td>
</tr>
</tbody>
</table>

Notes: Sample is U.S. House general election candidates from two major parties in 1956–2014 and the unit of observation is candidate by year. I further restrict the sample to elections where the opponent’s primary’s top two contestants are of opposite gender. I use triangular kernel that weights points near the threshold more than those far from the threshold and use the preferred choice of polynomial order of 1, which gives the standard local linear RD point estimator. The bandwidth are computed following the procedure by Calonico et al. (2017), which trade-off bias and variance of the treatment effect to find the MSE-optimal bandwidth. Column 6 has a smaller sample size because some candidates come from a single candidate primary. Standard errors clustered at the district level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Appendix B

This section derives the impact on the unconditional outcome. Impact on the unconditional outcome is estimated assuming that everyone has a defined outcome. With no defier assumption, unconditional outcome is

\[
\text{Impact on the unconditional outcome} = \text{Impact on the outcome for AT} \times P_r(\text{AT}) + \text{Impact on the outcome for C} \times P_r(C) + \text{Impact on the outcome for NT} \times P_r(\text{NT})
\]

or:

\[
E[Y_1W_1 - Y_0W_0] = E[Y_1 - Y_0|\text{AT}]Pr(\text{AT}) + E[Y_1 - Y_0|\text{C}]Pr(\text{C}) + E[Y_1 - Y_0|\text{NT}]Pr(\text{NT})
\]

where \(Y_1\) is the outcome conditional on winning when \(D = 1\) (candidate faces women), \(Y_0\) is the outcome conditional on winning when \(D = 0\) (candidate faces men), and \(W_0\) and \(W_1\) are indicator variables of winning the election when \(D = 0\) and \(D = 1\), respectively. We observe the outcome of candidates who win but do not observe the outcome of candidates who lose. Let’s assign an arbitrary value \(A\) as the outcome of the losers. Then we can re-write the above equation as:

\[
E[Y_1W_1 - Y_0W_0] = E[Y_1 - Y_0|\text{AT}]Pr(\text{AT}) + E[Y_1 - A|\text{C}]Pr(\text{C}) + E[A - A|\text{NT}]Pr(\text{NT})
\]

Formally define AT and C to write

\[
E[Y_1W_1 - Y_0W_0] = E[Y_1 - Y_0|\text{AT}]Pr(\text{AT}) + E[Y_1 - A|\text{C}]Pr(\text{C})
\]

\[
= E[Y_1 - Y_0|\text{AT}]Pr(\text{AT}) + E[Y_1 - A|\text{C}]Pr(\text{C})
\]

\[
= E[Y_1W_1 - Y_0W_0|W_1 = 1, W_0 = 1]Pr(W_1 = 1, W_0 = 1) + E[Y_1 - A|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)
\]

13. \(x = 0\) condition omitted for simplicity
Now add and subtract $E[Y_0|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)$ from equation (15)

$$E[Y_1W_1 - Y_0W_0] = E[Y_1 - Y_0|W_1 = 1, W_0 = 1]Pr(W_1 = 1, W_0 = 1)$$

$$+ E[Y_1 - A|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)$$

$$+ E[Y_0|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)$$

$$- E[Y_0|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)$$

Rearrange to get

$$E[Y_1W_1 - Y_0W_0] = E[Y_1 - Y_0|W_1 = 1]Pr(W_1 = 1)$$

$$+ E[Y_0|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)$$

$$- E[A|W_1 = 1, W_0 = 0]Pr(W_1 = 1, W_0 = 0)$$

This can be re-written as

$$E[Y_1W_1 - Y_0W_0] = E[Y_1 - Y_0|W_1 = 1]E[W_1] + E[Y_0 - A|W_1 > W_0 = 0]Pr(W_1 > W_0)$$

In plain English, impact on the unconditional ideology now can be written as

**Impact on unconditional ideology** = Impact on conditional ideology $\times$ Fraction of candidates who win when opposed to women $+$ (Hypothesis on ideology that compliers who didn’t win would have had if they won - arbitrary value we give them when they lose) $\times$ Impact on winning.

Plugging in back the $x = 0$ condition, I formally define impact on the unconditional outcome as

$$E[Y_1W_1 - Y_0W_0|x = 0] = E[Y_1 - Y_0|x = 0, W_1 = 1] \cdot E[W_1|x = 0]$$

$$+ E[Y_0|W_1 > W_0, x = 0] \cdot Pr(W_1 > W_0|x = 0)$$

14. It’s easily confirmed that rearranging equation (9) also gives us the same result
Now as a thought experiment, imagine plugging in different numbers for $A$. Plug in 0, -2, 10, 500... The left hand side of the equation (19) will also change as we plug in different numbers for $A$. The higher the number we choose for $A$, the smaller the RD effect on the outcome will be, and the lower the number we choose for $A$, the higher the RD effect on the outcome will be. Now rearrange to write in terms of conditional outcome:

$$
E[Y_1 - Y_0|x = 0, W_1 = 1] = \frac{1}{E[W_1|x = 0]} \cdot \left( E[Y_1W_1 - Y_0W_0|x = 0] \right)
$$

$$
- \underbrace{Prob(W_1 > W_0|x = 0)}_{\text{RD effect on winning}} \cdot \underbrace{E[Y_0|x = 0, W_1 > W_0]}_{\text{Unobservable}}
+ \underbrace{Prob(W_1 > W_0|x = 0)}_{\text{RD effect on winning}} \cdot \underbrace{E[A|x = 0, W_1 > W_0]}_{\text{Arbitrary number}}
$$

(18)

We saw from equation (15) that whatever number we choose for $A$, this change gets reflected also on the impact on the unconditional outcome. Equation (20) shows that those two offset each other when estimating the impact on the unconditional outcome. I simulate bounding strategy for the impact on the unconditional outcome, using different values of $A$. The result is shown in Table B.
Table B. Impact on the conditional outcome, with different arbitrary values

<table>
<thead>
<tr>
<th>Arbitrary value</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower bound</td>
<td>0.190**</td>
<td>0.184**</td>
<td>0.174*</td>
<td>0.179*</td>
<td>0.178*</td>
</tr>
<tr>
<td>Boot. std error</td>
<td>0.091</td>
<td>0.091</td>
<td>0.096</td>
<td>0.097</td>
<td>0.097</td>
</tr>
<tr>
<td>Upper bound</td>
<td>0.451***</td>
<td>0.445***</td>
<td>0.435***</td>
<td>0.440***</td>
<td>0.440***</td>
</tr>
<tr>
<td>Boot. std error</td>
<td>0.161</td>
<td>0.160</td>
<td>0.165</td>
<td>0.161</td>
<td>0.161</td>
</tr>
</tbody>
</table>

Notes: Table B shows that different arbitrary values I assign to the outcome of the losers of the race do not change bounds on the conditional outcome by much. All columns assume that compliers who faced men would have assumed the average political score of the compliers who faced women and thus won the race.