ISSN 2279-9362

The Impact of the 2017 Women's March on Female Political Representation

Alessandra Moresi

No. 669
February 2022

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

The Impact of the 2017 Women's March on Female Political Representation*

Alessandra Moresi†

February 23^{rd}, 2022

Abstract

The 2017 Women's March has been the largest single-day protest that ever happened in the US and a showcase of female leadership. This paper exploits the geographic variation in congressional districts' exposure to protests to investigate whether the March had an impact on the supply of female politicians in partisan primaries. Using a difference-in-differences design and an event-study analysis that leverage distance from the nearest protest, I find that the March caused a marked increase of the probability of having at least one female candidate (59%) and of the share of female candidates (58%) in Republican primary races. I find smaller effects in Democratic primaries. Moreover, I investigate the consequences for women's representation in federal politics, finding that the March increased the probability of having a female House Representative by 61%, regardless of party affiliation.

JEL Codes:

Keywords: protest; identity; gender; politics

*I am grateful to Giovanni Mastrobuoni, Juan Morales and Daniela Del Boca for invaluable guidance and support. I also acknowledge constructive feedback and discussions from participants of the 2021 Researchers' Forum at Collegio Carlo Alberto and University of Turin. All errors remain my own.

†University of Turin and Collegio Carlo Alberto.

Contact email: alessandra.moresi@carloalberto.org Phone: +39 3208260738

1 Introduction

Women make up one fourth of parliamentarians worldwide, a number that is far below equal political representation (UN, 2020). As of 2020, women accounted for 27% of seats in the US House of Representatives, a figure that has steadily increased from 5% in 1990 (CAWP 2020). Despite the progress made during the last 30 years, the persistence of this gender disparity may suggest that women face barriers in accessing public offices at the federal level. Top-down policies such as the introduction of gender quotas have been shown to be effective in reducing the gender gap in political representation (Beaman et al., 2009; Pande and Ford, 2012; Casas-Arce and Saiz, 2015), but little is know about how bottom-up feminist activism may ease the constraints faced by women and improve female political representation.

Barriers to equal political representation could consist of either voters or parties being biased in favor of male politicians. Nonetheless, in the American context women seem not to face voters' discrimination (Anastasopoulos, 2016), while parties' bias against female politicians conflicts with their desire to maximise electoral outcomes (Casas-Arce and Saiz, 2015). Another barrier faced by females who want to enter into federal politics could be that they face higher societal costs when acting as political leaders, for instance because of patriarchal social norms (Akerlof and Kranton, 2000). Unequal female political representation could therefore stem from women's differential willingness to run for office: if females do not want to act as political leaders, then we can only observe male dominated parliaments.

There could be competing explanations to why women do not want to act as political leaders: Wasserman (Forthcoming) shows the existence of a gender gap in women's propensity to run for local offices after an electoral loss, suggesting that the psychological burden of failure differs between males and females and establishing that previous experience in politics has a mitigating effect (Wasserman, 2021). On the other hand, negative attitudes towards female leaders could explain women's differential willingness to run for office: if the society does not expect women to behave as leaders, breaking the norm and running for office entails bearing a cost.

Can feminist mobilization shape the social perception of female leaders up to the point of affecting women's choice to run for office? And what are the consequences for female political representation? To provide evidence on this issue, I study the 2017 Women's March: the non violent protest that inaugurated

Donald Trump's presidential mandate. This case is a well-suited testing ground for at least three reasons: first, the 2017 Women's March was the largest singe-day protest with both a political and a feminist root that ever happened in the US, with a turnout of more than 3 million people; second, the March was an unprecedented showcase of female leadership (see Figure 5); third, the Women's March on the US Capitol was accompanied by 613 sister Marches, providing enough geographic variation to allow for identification.

The main empirical challenge in estimating the impact of the March is that historical ties with feminist movements and local dissent for the Trump administration are determinants of protests that also correlate with women's willingness to run for office and with their probability of being elected in the US House. Another challenge is reverse causality: it is ex-ante unclear whether it is the March that has an impact on female politicians or the other way around. To address these problems, I exploit the geographical partition of the US electoral process to build a geographical measure of congressional districts' exposure to protests. The distance between a district's electorate bulk and its nearest protest constitutes a continuous treatment variable that, once discretized, provides enough exogenous variation to both assess the impact of the March in a difference in differences design and to provide extensive evidence in support of the parallel trend assumption in an event study analysis. More specifically, I construct population-weighted and geographical protest buffers that allow to classify congressional districts as *close* (i.e. treated) and *not close* (i.e. untreated) in order to assess the causal impact of the 2017 Women's March. Treatment status is defined over proximity to protests because I want to measure the impact of exposure to local political and feminist activism (Fisher, 2019).

I use historical records of US House partisan primaries for both Republican and Democratic races to analyse the responsiveness of the extensive and intensive margin of the supply of female politicians to the March. Partisan primaries make an ideal setting for testing whether the March affected women's choice to run for office because they are the ground that parties use to select who shall be listed on the general election ticket. I use presence of at least one female candidate in the partisan primary to measure the extensive margin of the supply and the share of women running for office to measure the intensive margin. While the extensive margin of the supply matters for improving female representation in federal politics, the intensive margin is more likely to be relatively detached from parties' strategies. I find that

being *close* to the March increased the probability of having at least one female candidate in primary races: by 59% in Republican primaries and by 33% in Democratic races. The feminist protest also had an impact on the share of female candidates in Republican races, causing a 58% increase, while I find no evidence of such an effect within Democratic races. Results are qualitatively and quantitatively stable across 7 different population-weighted and 2 geographical protest buffers. Within those buffers, I also empirically assess the lack of pre-treatment differential time trends and I find the parallel trends to hold pre-treatment.

To asses the implication of the change in women's willingness to run for office on female political representation, I use general election returns. More specifically, I focus on the probability to elect a female US House representative and I find an increase of 61% with respect to the sample pre-treatment probability. This result holds across 6 different population-weighted and 3 geographical protest buffers, reinforcing the ground for a causal interpretation of the identified effect.

Results seem to suggest that the March shifted attitudes towards female leaders up to the point of affecting women's choice to run for office and female political representation at the federal level. There could be competing explanations to why the effects are stronger among Republicans than among Democrats. On one hand, it seems reasonable to think that the constraints imposed by negative attitudes towards female political leaders are more binding among conservatives, which implies that the supply of female Republican politicians is more responsive to reductions in the social cost of acting as leaders. On the other hand, it could also be that the complementarity between the March and the Democratic Party electoral campaigning strategy prevents identification within Democratic races (Fisher, 2019). Last, it could be that the Republican Party favored females' candidatures in primaries *close* to the March because it wanted a woman to list on the general election ballot, so as to maximise electoral gains by means of *strategic descriptive representation* (Weeks et al., Forthcoming). This latter hypothesis seems to be confirmed by the historical evolution of the supply of female candidates in Republican primaries (Figures 2 and 3): we don't observe a jump of female candidatures in Republican primaries after the March, which suggests that Republican women placed their candidatures in *close* districts with the aim of capitalizing on the March.

This paper speaks to several strands of the literature, encompassing the fields of economics, sociol-

ogy and political science. The theoretical roots of the empirical analysis lie in identity economics, the theory that posits how social movements transform social categories and norms (Akerlof and Kranton, 2000; Kranton, 2016). The empirical study that most closely relates to this paper is Levy and Mattsson (2021), who find that the #MeToo movement influenced attitudes towards gender violence up to the point of influencing women's choice to report sexual harassment crimes. As regards the comparability of my findings with the literature, in a contemporaneous paper Larreboure and González (2021) analyse the impact of the 2017 Women's March on the votes obtained by women and ethnic minorities in the 2018 US House general elections. Using a weather shock as an instrument for county-level protest size, they find that 1000 extra protesters increased the vote shares for women by 32% with respect to the sample mean. In an ancillary analysis (reported in Appendix C), I apply my identification strategy to their same outcome and I show that being *close* to the March increased the vote shares for women in general elections by 37% of the sample mean. Furthermore, I contribute to the debate on the impact of the 2017 Women's March in at least three ways: first, my unit of analysis are congressional districts rather than counties; second, I apply a difference in differences identification strategy rather than an instrumental variables approach; third, I analyse the supply of female politicians in partisan primaries rather than the demand for female politicians in general elections.

This study also contributes to the empirical works analysing the impact of protests on attitudes, beliefs and policy preferences in the United States, establishing that female-led protests affect the social perception of female leaders up to the point of affecting the behavior of female politicians. Madestam et al. (2013) studies the 2009 Tea Party rallies and show that, among other outcomes, the event shifted public opinion towards the views expressed by the movements' leaders. A large number of studies analyze the impact of the Black Lives Matter (BLM) movement: Mazumder (2019) studies how the 2014 wave of BLM protests had an impact on racial attitudes, weakening racial prejudice among the youngest cohorts but strengthening it among the oldest; Reny and Newman (2021) use the random timing of George Floyd's assassination to show how the 2020 wave of BLM affected public opinion, fuelling negative attitudes towards the police. Similarly, Ebbinghaus, Bailey and Rubel (2021) shows how BLM led to the adoption of state police reforms, while Klein Teeselink and Melios (2021) use a rainfall shock to assess the impact of BLM on the 2020 presidential elections. Wasow (2020) analyzes how the 1960s black led

protests affected public opinion, showing the effectiveness of non violent protests in shifting individuals' beliefs. Adopting an historical perspective Mazumder (2018) describes how non violent civil rights movements had the power to prime the American identity, ultimately changing individuals' beliefs and shaping the political landscape even years after their occurrence.

Last, my findings also relate to the literature on the impact of the Arab Spring uprisings (a female-led wave of protests) on women's empowerment (Bargain, Boutin and Champeaux, 2019; El-Mallakh, Maurel and Speciale, 2018), and support the studies claiming that exposure to strong female role models has an impact on women's behavior (see, among others, Matsa and Miller, 2011; Jensen and Oster, 2009; La Ferrara, Chong and Duryea, 2012). The remainder of the paper is organized as follows: section 2 contextualizes the 2017 Women's March and details the institutional context of federal elections; section 3 describes the data used throughout the analysis; section 4 explains the empirical strategy, section 6 the robustness checks and section 7 concludes.

2 Background

2.1 The 2017 Women's March

The day of Donald Trump's presidential victory, an elder lady from Hawaii posted on a Facebook page a simple statement: "$I\ think\ we\ should\ march.$" (Kearney, 2016). Her post went viral overnight and spurred the largest single-day protest in the US history to-date: 1% of US residents took the streets a few weeks later, on January 21st, 2017 (Broomflied, 2017; Easley, 2017).

American citizens marched because they were outraged by Trump's misogynist declarations (Lusher, 2016) and concerned that his presidency would threaten civil rights, the environment, reproductive rights and, most importantly, the right to equality. The success of the movement is related to the two principles of *intersectionality* and *inclusion*: diverse interests groups, motivated by concerns about gender, race and class, gathered behind the women's flag, advocating change (Fisher, Dow and Ray, 2017). As a result, the movement was an intersectional coalition promoting equity, a structural difference with respect to previous feminist movements.[1]. Despite the diverse interest groups that supported the March, the event

[1] *Intersectionality* is a characteristic of the fourth wave of American feminism (Fisher, Dow and Ray, 2017). Among the organizers of the 2017 Women's March we find: *Black Lives Matters, Occupy Wall Street, OkayAfrica, Gatherings for*

was a manifestation of female leadership: millions of people wearing *pussyhats* (i.e. pink hats knitted by volunteers) took to the streets and showed the world what women were capable of organizing (Pussyhat Project n.d., see also Picture 5).

The March had a twofold legacy: it inaugurated a series of protests against the Trump administration and it favored the emergence of local *resistance* groups.[2] Fisher (2019) describes how the same *resisters* that marched on January 21^{st} kept on marching on DC up to the 2018 midterm elections, while at the same time they actively engaged in electoral campaigning for Democratic candidates. The *resisters* directly targeted the districts that needed it most, with the explicit intention of *"Flipping seats from red to blue and winning democratic majorities."* (Swing Left, n.d.).[3]

Coherently with the fact that it was a mobilization against the newly elected president, the largest Women's March (in absolute terms) took place in Washington DC (Jamieson, 2016).[4] Nonetheless, several sister Marches were organized both within and outside the country. Most notably, US citizens gathered in symbolic places to manifest their outrage: map 4 shows the share of population turning out to protest in populated places and suggests that the largest March (in relative terms) took place in the town of Seneca Falls, highlighting the historical ties of the March with previous feminist movements.[5] Since most *resisters* gathered in symbolic places and then went back home to engage in local political activities, protest size is likely to be affected by measurement error.[6]

While protest size is unlikely to capture the local strength of *resistance* groups, geographical proximity to protests could rather be a valid exposure variable: the distance from any march, regardless of its size, is a valid proxy for the constituency's exposure to local *resistance* activity. Indeed, it seems

Justice and the *American Civil Liberties Union*, while pro-life movements were excluded from the protests (Fisher, 2019). Most importantly, American feminism has been historically criticized for having a mostly white liberal identity (Davis, 1981) Interestingly, Larreboure and González (2021) find no evidence of an effect of the March on the vote shares for women from ethnic minorities, which is coherent with the long-standing critiques about American feminism.

[2]Among which we can name the following: March for Science, People Climate March, March for Racial Justice, March for our Lives and the Families Belong Together (Fisher, 2019).

[3]As an additional anecdote, Amanda Arias Lewis (the executive assistant of the Women March organization), was the office manager of senator Sander's private suite for the 2020 presidential elections (Fisher, 2019).

[4]500,000 people took to the streets in the federal capital.

[5]The first American feminist wave was inaugurated by the Seneca Falls convention in 1848. The meeting anticipated the debate on women's right to vote in elections by more than two decades.

[6]In a preliminary analysis I use three different definitions of protest size as continuous treatment variables and I find no evidence of an effect of the March. This feature is likely due to classical measurement error in protest size, which biases estimates towards zero. I also investigate the responsiveness of the outcome variables when using the distance as continuous treatment variable. All the results reinforce the findings of the binary difference in differences and are available upon request. When using size, distance and their interaction as treatment variables, I find that the marginal effect of size is always small (virtually zero) and rarely significant.

reasonable to think that a sister March occurs if and only if there is a cluster of *resisters* who organize it. While it is hard to believe that geographical proximity to protests is as good as randomly assigned, the data generating process of the distance between the congressional district electorate's bulk and the nearest protest provides enough exogenous variation to allow for the identification of the causal effect of the March on women's choice to run for office and on the probability that they are elected as US House representatives. Furthermore, the extensive evidence in support of the lack of lead effects of the March discussed in section 5.2 should reassure the reader in that the estimated parameters are not biased by reverse causality nor omitted variables.

2.2 Institutional Context

The House of Representatives is the lower chamber of the bicameral legislature of the US federal government and it is composed by 435 voting members. Each representative is elected to represent one congressional district for a two-year mandate through single member plurality voting (US Gov, 2020).[7]. Congressional districts (CD or districts from now on) are apportioned to states on a population basis. Every 10 years (i.e. 5 elections) the Federal Government re-apportions CD to states based on the decennial census (US Census, 2020). Most state parliaments have the authority to re-draw CD boundaries before elections (Ballotpedia, 2020*b*).[8] As a result of this institutional framework, districts are not stable geographical units but rather an administrative partitioning of the federal electoral process.

To run for office as US House representative, an individual must satisfy three federal requirements: having attained the age of twenty-five years, being a citizen of the US since at least seven years and being an inhabitant of the state where s/he is elected (US Constitution, 2020).[9] On top of federal requirements, candidates running for office must satisfy a variety of state-level provisions. Indeed, the US is one of the few countries that does not have federal ballot access laws (Ballotpedia, 2020*a*).

[7] In SMP electoral systems voters cast a ballot directly for candidates rather than for parties. SMP voting system is also known with the name of *first past the post* SMP voting systems favour the party that can influence the distribution of its voters across electoral districts. Indeed, winning by small margins in most of districts and losing by large margins in all the others can still pay with a representation premium.

[8] 33/50 states parliaments can rule over redistricting by means of a state law. The law is introduced by the party who rules the state legislature (state's upper and lower chamber). The remainder states rely on different redistricting procedures.

[9] The only federal requirement that is disciplined outside the US Constitution concerns the financing of electoral campaigns. Candidates must file a petition to the Federal Election Commission (FEC) within 15 days from making campaign expenditures exceeding $5,000. The FEC cannot prevent any individual who correctly files the petition from running for office and has to authorize the candidates within 10 days from receiving the relevant documentation (FEC, 2020).

Ballot access laws usually impose two requirements to congressional candidates: to file a petition with a minimum number of signatures from citizens endorsing the candidature and to pay a registration fee.[10] The minimum required number of signatures, the registration fee and the deadlines for complying with the filing requirements may vary from one state to another and from one election to the next (Ballotpedia, 2020*a*). Within this regulatory framework, a US citizen has three different ways to ensure that s/he is entitled to a federal office: s/he can seek the nomination of a state recognized political party, s/he can run as an independent candidate and s/he can run as a write-in candidate.

Primary elections constitute the ideal setting to test whether the March affected women's choice to run for office because they are the ground used by political parties to nominate the candidates they want to list on the general election ballot. Any citizen who meets both federal and state level ballot access requirements has the right to compete in the primary. In the US, primary elections are disciplined by state laws which are often nuanced. This results in the absence of a normative listing of the types of primaries. Furthermore, since primaries are disciplined by state laws, also their legal framework can vary from one election to the next. The National Conference of the State Legislatures classifies primary elections in seven different categories based on their openness: closed, partially closed, partially open, open to unaffiliated voters, open, top two and top four (NCSL, 2020). This categorization emphasizes that primary election systems usually differ in the pool of voters entitled to cast a ballot.

Most notably, while general elections regularly take place every two years, the regulatory fragmentation that characterizes primary elections implies that we do not observe Republican nor Democratic primary races for each House seat that is up for election.[11] Nonetheless, both the Republican Party (also known as Great Old Party, i.e. GOP) and the Democratic Party (DP) usually call the primaries before general elections. However, in the US 90% of Congress members gets re-elected (Fowler and Hall, 2015), which translates in most primary races being de facto uncontested.

While understanding the interplay between the state-level laws that discipline primaries and party dynamics is beyond the scope of this project, it is important to acknowledge the normative fragmentation

[10] In some states both the petition and the fee are normative requirements, in others it is sufficient to meet one of them. In most of the states the petition is an essential requirement.

[11] General elections take place after primaries, and they coincide with midterm and presidential elections. General elections usually determine who will seat in the House for the following two years. There are rare cases when a candidate running in a top-two or top-four primary system gets the absolute majority of votes. When this happens, there is no need to call the general elections and the primary winner is directly entitled to seat in the House. Top-two and top-four partisan primary races are excluded from my analysis.

that regulates the federal electoral process and to adopt the appropriate coping strategies: in the empirical strategy (section 4), I consider the baseline DiD and event study specifications the ones that include state-election fixed effects. Indeed, to be sure that the identified effects are not driven by changes in ballot access or primary laws, I exploit the variation in geographical proximity to protests within an election year.

3 Data

To empirically assess the impact of the March on women's choice to run for office and its consequences for female political representation, I build a novel CD level dataset using several data sources. The main challenge in the construction of a CD level panel is redistricting. To address the changes in electoral geography that may plague my estimates, I follow Fowler and Hall (2015) and code a district as new every time that redistricting occurs: if a state passes a redistricting bill, then all the CD within that state are coded with a new identifier. As a result, I obtain an unbalanced CD level panel for the years 2012-2020. Redistricting information comes from Miller and Camberg (2020) and Jacobson (2015) and is available up to 2018.[12] Across the 2012-2018 elections, four states were affected by redistricting: Florida, North Carolina and Virginia in 2016 and Pennsylvania in 2018. The Pennsylvanian 2018 congressional redistricting was mandated by the state's Supreme Court with the aim of remedying to an illegal partisan gerrymander.[13] Since Pennsylvania's districts are coded with a new identifier from 2018, these CD do not contribute to the identification of the causal effect of the March. Moreover, Florida, North Carolina and Virginia had stable CD boundaries across the 2016 and 2018 elections, which implies that these states contribute to the estimation of the causal effect of the March. Nonetheless, CD are classified as *close* (i.e. treated) and *not close* (i.e. not treated) based on the time-invariant distance from the nearest protest: if a state had passed a redistricting bill, then the lagged distance becomes meaningless. To cope with this potential caveat, the states affected by redistricting in 2016 are excluded from the event study analysis.[14]

[12] Redistricting information on the 2020 election cycle is available in the America Votes 34 (2021) and will be incorporated in the analysis as soon as the Miller and Camberg (2020) data for the 2020 election cycle will be released.

[13] Pennsylvanian GOP representatives petitioned the Supreme Court to oppose the redistricting plan, but failed.

[14] In particular, I drop the rows referring to Florida, North Carolina and Virginia for the years 2012 and 2014 when estimating equations 4 and 5. Note that the results are robust to using a balanced panel of administrative units and including state-election fixed effects to capture changes in political geographies. I also plan to follow the recent political economy literature (e.g. Autor et al. 2020; Calderon, Fouka and Tabellini 2021) and use population crosswalk files to bridge political geographies to further

I begin by collecting historical records of US House Democratic and Republican primaries (Miller and Camberg, 2020).[15] I focus on two main variables: a dummy for having at least one female candidate in the primary race and share of women running in each party's primary. These measures allow me to analyse the impact of the March on the extensive and intensive margin of the supply of female politicians. To analyse the impact of the March on the probability that a woman is elected in general elections, I augment the data with records from the Center for American Women and Politics (CAWP, 2020) and code a dummy for female US House representative.

I also use protest data from the Crowd Counting Consortium (CCC; Chenoweth and Pressman, 2017; Fisher et al., 2019). The CCC is an academic project whose mission is to collect information on marches, protests, strikes and riots in the US. Data are crowd sourced and verified by affiliated scholars through fact-checking. Sobolev et al. (2020) show the reliability of the CCC data as compared to protest estimates from cell phones and Twitter data. The raw crowd estimates contain information on any protest event, including minor marches which are unlikely to be a good proxy for *resistance* activity.[16] According to the CCC, there were 614 Women's Marches throughout the US. To infer the exact geographical coordinates of each March, I follow Wallace, Zepeda-Millán and Jones-Correa (2014) and rely on the GIS geocoding service.[17] Next, to address the crowd-sourced nature of the CCC data, I aggregate the protests at the urban level thanks to urban areas TIGER/Line shapefiles (US Census Bureau, 2020*b*).[18] This procedure allows me to identify 412 major protests happening in cities which are a good proxy for *resistance* activity.

Last, I focus on the geographical partition of the US electoral process in congressional districts as of 2018 (the first election year after the March). I retrieve 2018 CD's population-weighted centroids from the Geographic Correspondence Engine of the Missouri Census Data Center (2020) and I compute 2018 CD's geographical centroids using CD TIGER/Line shapefiles (US Census Bureau, 2020*c*). I use population-weighted centroids to proxy the geolocation of the electorate's bulk in each district and geographical

assess the validity of my estimates. I plan to build the CD level panel borrowing from the recent work of Ferrara et al. (2021).

[15] Primary election returns are available only up to 2018. Primary election returns for the 2020 election cycle are available in the America Votes 34 (2021) (released in December 2021). Miller and Camberg plan to work on the digitalization of such returns during the forthcoming summer. The Miller and Camberg (2020) data do not contain primary election returns for the states that adopt a top-two or top-four primary system (i.e. California and Washington) and for Louisiana, where partisan primaries do not take place.

[16] There are multiple records of Marches with only a few people turning out to protest.

[17] Before resorting to the GIS geocoding service, I manually aggregate protest locations using the US cities comprehensive database (Simplemaps, 2020) to make sure that all the protest locations correspond to a populated place listed in the US Census Bureau. This procedure leaves me with 604 protest locations, which I then classify as urban/non urban.

[18] More specifically, I perform a spatial join between geolocated protests and urban areas in GIS.

centroids to test the robustness of my results. The distance between CD's centroids and the border of the closest urban polygon hosting a March is the measure of constituents' exposure to treatment on which I build the empirical strategy.[19]

Time varying control variables such as the share of votes obtained by the Democratic Party in the previous House election and population density come from the Dave Leip's Election Atlas (2020) and the US Census Bureau (2020*a*). Table 1 reports summary statistics of all the described variables. Table 2 states the definition of distance variables, while Tables 6 and A1 show the distribution of population-weighted and of geographical distance. Figure B3 shows the density of population-weighted distance after partialling out state fixed effects.

4 Empirical Strategy

This paper relies on the discretization of the continuous distance variable to assess the differential impact of being close to the March in a difference in differences and in an event study framework. I test the hypotheses that being *close* to the March led women to enter into federal politics, ultimately changing the results of general elections. In this section, I detail the empirical strategy when defining as *close* all CD located within the 75^{th} percentile of the population-weighted distance variable (i.e. the distance between the 2018 CD population-weighted centroids and the border of the closest urban polygon hosting a March), which corresponds to a protest buffer of 28.5 km.

I estimate the following difference in differences equation, separately for Democratic and Republican primary races:

$$y_{pdt} = \theta_d + \Gamma(d)_{st} + \delta POST_t \cdot close_d + X'_{dt}\mu + \epsilon_{pdt} \tag{1}$$

Where y_{pdt} is a dummy for having at least one female candidate and the share of female candidates in the partisan primary p of district d in election year t. θ_d are CD fixed effects and $\Gamma(d)_{st}$ are state-election fixed effects that control for changes in ballot access and primary laws that might change women's incentives to run for office.[20] I also assess the stability of the DiD coefficient (δ) when allowing changes

[19] See Figure C7 for a visualization of the two layers of geography that allow to compute the population-weighted distance

[20] Note that CD are geographically nested in states, while election years and primary type fixed effects are nested in state-election fixed effects.

in laws to be captured by state specific linear time trends rather than by state-election fixed effects.[21] $POST$ is a dummy for the election years after the March (i.e. 2018 and 2020) and $close_d$ is a dummy for all CD whose population-weighted centroid falls within the 75^{th} percentile of the distance variable. X'_{dt} are time varying controls: population density and the share of votes obtained by the Democratic Party in the previous US House elections.

As regards the probability that a woman is elected to represent the district in the US House, I estimate the following equation:

$$y_{dt} = \theta_d + \Gamma(d)_{st} + \alpha POST_t \cdot close_d + X'_{dt}\mu + \epsilon_{dt} \tag{3}$$

Where y_{dt} is a dummy for female US House representative and the remaining components are defined as in equation (1). The only difference between equation (1) and (3) is that general election returns do not have a partisan dimension (p).

The main challenge to the identification of the causal effect of the March is that protests are more likely to occur (and to be larger in magnitude) in places with greater underlying political grievance and historical ties with the feminist movement (see map 4). Furthermore, Sobolev et al. (2020) show that rallies were larger in more densely populated and democratic places.[22] While I can directly control for population density and for the share of votes obtained by the DP in the previous US House elections, local dissent for the Trump administration and historical ties with feminist movements are unobservable determinants of protests that are likely to correlate also with the supply of female politicians and with the probability that a woman is elected to represent the district. It is therefore crucial to establish the plausibility of the identifying assumption in equations (1) and (3) to give a causal interpretation to the DiD coefficient.

Difference in differences provide unbiased estimates of the parameter of interest (δ in eq (1), α in eq (3)) under the assumption that the counterfactual trend behaviour of the treated (i.e. $close$) districts would have been the same as the one of control districts (i.e. $not\ close$) in the absence of treatment. A concern

[21] More specifically, I estimate the following equation:

$$y_{pdt} = \theta_d + \Phi_s \cdot trend_t + \delta POST_t \cdot close_d + X'_{dt}\mu + \epsilon_{pdt} \tag{2}$$

[22] Such analysis was present in the working paper version of the published paper.

to identification would arise if $close$ and $not\ close$ districts were characterized by different post treatment growth rates in the potential outcomes. For instance, if we think that districts $close$ to the protests are at the same time affected by changes in ballot access or primary laws that change women's incentives to run for office, it might be hard to believe that the parallel trend assumption would continue to hold post treatment. Four things should be reassuring in this regard: first, I control for changes in ballot access provisions and primary laws in two different ways and the estimated parameters are highly comparable across model specifications; second, the distance variable that I discretize arises from a geostatistical analysis of the administrative partition of the electoral process and is likely to be relatively exogenous to district specific differential time trends; third, I formally test the lack of pre-treatment differential time trends through an event-study analysis; fourth, the lack of pre-treatment differential time trends holds across 7 different population-weighted protest buffers and 2 geographical protest buffers (see section 6).

I formally falsify the parallel trend assumption for the supply of female politicians by estimating equation (4) separately for Democratic and Republican primaries:

$$y_{pdt} = \theta_d + \Gamma(d)_{st} + \sum_{\substack{\tau=2012 \\ with\ \tau \neq 2016}}^{\tau=2018} \delta_\tau \cdot close_{d\tau} + X'_{dt}\mu + \epsilon_{pdt} \tag{4}$$

Equation (4) has the same structure as equation (1), with the only difference being that $close_{d\tau}$ is a time-varying battery of dummies for $close$ districts in each election year (τ). Hence, δ_τ are the estimated parameters for the lagged variables and for the lead variable. The absence of lead effects in estimates of equation (4) would be suggestive of the absence of violation of the parallel trend assumption.

Last, I estimate equation (5) to falsify the parallel trend assumption for the probability that a woman is elected to represent the district in general elections:

$$y_{dt} = \theta_d + \Gamma(d)_{st} + \sum_{\substack{\tau=2012 \\ with\ \tau \neq 2016}}^{\tau=2020} \alpha_\tau \cdot close_{d\tau} + X'_{dt}\mu + \epsilon_{dt} \tag{5}$$

Here, again, the only difference between equation (4) and (5) is that general elections do not have a partisan dimension (p).

5 Results

5.1 Difference in differences

Table 3 and 4 present the estimates of the causal effect of being *close* to the March on the extensive and intensive margin of the supply of female candidates in partisan primaries (equation (1)). Columns 1-3 report the estimates on Republican races, columns 4-6 on Democratic races. Columns 1 and 4 show the baseline difference in difference specification, columns 2 and 5 add the time-varying controls (population density and support for the DP in the previous US House election) and columns 3 and 6 substitute state-election fixed effects with state-specific linear time trends. Throughout the analysis, standard errors are clustered at the CD level.

Across both partisan primaries and model specifications reported in Table 3, the point estimate of the DiD coefficient (δ in equation (1)) for the probability of having at least one female candidate in the party's primary is stable around 13 percentage points and statistically significant at conventional levels. The adjusted r-squared decreases when introducing control variables, suggesting that the penalty term outweighs the predictive power of the time-varying CD characteristics in explaining the variability of the dependent variable (i.e. time-varying controls seem not to matter given the CD fixed effects).[23] Most notably, the relative magnitude of the estimated effects greatly differs across Republican and Democratic primaries: GOP races *close* to the March experienced an increase in the probability of 59% with respect to the sample pre-treatment probability, while for DP races the increase is 33% only.[24]

Table 4 reports estimates of equation (1) for the shares of female candidates in each party's primary. While the March caused an increase of 58% in the share of female candidates running in Republican races, I find no evidence of such an effect among Democratic races.

As regards the effect of the March on general election results, Table 5 presents estimates of equation (3) for the probability of electing a female US House representative. Columns 1-3 pool together Republican and Democratic female representatives, columns 4-6 refer to the probability of electing a female

[23] Given the behavior of the adjusted r-squared, I do not include additional time-varying controls to avoid adding noise.

[24]
$$\Delta P(woman\ candidate) = \frac{\delta}{Dep\ var.\ mean} \tag{6}$$

GOP representative and columns 7-9 focus on female DP representatives. Similarly to Tables 3 and 4, columns 1, 4 and 7 report the baseline DiD specification, columnn 2, 5 and 8 add time varying controls and columns 3, 6 and 9 substitute state-election fixed effects with state-specific linear time trends. The point estimate of the DiD coefficient (α in equation (3)) for the probability of electing a female US House representative -regardless of party affiliation- is 11 percentage points, which corresponds to an increase of 61% of the sample pre-treatment probability. Results seem to be driven by Democratic women, for which I find an increase of 64% of the sample pre-treatment probability, while I find no evidence of an effect for Republican women. This result is coherent with the demand for progressive female politicians brought about by the *resisters*. In Appendix C, I also analyse the impact of being *close* to the March on the share of votes obtained by women in 2018 and 2020 general elections (see columns 1-3 of Table C2 and Figures C1a, C2a, C3a). I find that being *close* to the March caused an increase of 37% of the sample pre-treatment votes share. My estimate is highly comparable to Larreboure and González (2021), who find that 1000 extra protesters caused an increase of 32% of the sample mean.[25]

5.2 Event Study

Figures 6 and 7 plot the estimated parameters for the lead and lagged treatment variable, meaning δ_τ in equation (4).

As regards the effect of the March on the probability of having at least one female candidate in the GOP primary, the estimated effect of the March (i.e. $\delta_{\tau=2018}$, equation (4)) is 14 percentage points, which corresponds to an increase of 63% of the sample pre-treatment probability. The identified effect is also significant at conventional levels and highly comparable to the classical DiD specification (i.e. estimates of δ in equation (1)). When estimating the same equation on DP primaries, I find an increase of 39% of the pre-treatment probability which is barely non-significant at conventional levels (point estimate=0.14, C.I.=-0.03/0.30). Nonetheless, Figure 6 seems to suggest the absence of any lead effect (i.e $\delta_\tau = 0 \; \forall \; \tau < 2018$), which should reassure the reader about the plausibility of the parallel trend assumption.

Similarly, the event study estimates on the share of women running for office in the Republican

[25] I choose to relegate the analysis of the vote shares obtained by women in Appendix C because it does not constitute the core of my analysis.

primaries (Figure 7) suggest that the March caused a surge of 67% in the sample pre-treatment share of female candidates, which is also significant at conventional levels. As with the canonical DiD specification, I find no evidence of such an effect in the Democratic primaries. Estimates of the lead effects suggest that the parallel trend assumption is likely to be plausible also for this outcome (i.e. $\delta_\tau = 0 \; \forall \; \tau < 2018$).

Moreover, Figure 8 shows the estimated α_τ of equation (5), allowing the evaluation of the plausibility of the parallel trend assumption for the probability of electing a female US House representative. Panel (a) reports estimates for the pooled probability of electing a woman, regardless of party affiliation, while panels (b) and (c) respectively show the probability of electing a Republican or a Democratic woman. While the parallel trend assumption seems to hold in panels (a) and (b), it is hard to believe its plausibility for the probability of electing a Democratic female US House representative. The estimated causal effect of the March on the probability of electing any female representative is around 10 percentage points for both the 2018 and 2020 General Elections, which correspond to a 56% increase with respect to the sample pre-treatment probability.

6 Robustness Checks

A valid concern related to the distance-based DiD and event study estimates is whether the estimated parameters are identifiable only when using an arbitrary chosen protest buffer. To address this potential caveat, in this section I check the robustness of my results across different population-weighted and geographical protest buffers.[26]

6.1 Robustness across population-weighted protest buffers

The population-weighted distance variable is defined as the distance between the 2018 CD population-weighted centroids and the border of the nearest urban polygon hosting a protest (see Table 2). In section 4, I classify as *close* all the districts whose centroid falls within the 75^{th} percentile of the population-weighted distance, which corresponds to a population-weighted protest buffer of 28.5 km. In this section, I investigate what happens to the estimated parameters across different protest buffers.

[26] I also plan to work on further robustness. More specifically, I will assess the robustness of the inference when correcting standard errors for spatial correlation (Conley, 1999; Conley, Hansen and Rossi, 2012).

Table 6 shows the distribution of the population-weighted distance and the corresponding population-weighted protest buffers. It is straightforward to notice that 43% of population-weighted CD centroids fall within protesting urban area polygons and that the variable is overdispersed.

Based on the distribution of this variable, I begin by classifying as *close* all CD located within the 50^{th} centile of the population-weighted distance and then I progressively classify as *close* all CD located within the 55^{th}, 60^{th}, 65^{th}, 70^{th}, 75^{th}, 80^{th}, 90^{th}, 95^{th} and 100^{th} centiles. Then, I estimate a battery of difference in differences specifications (i.e. equations (1) and (3)), changing each time the definition of *close*. Figures 9, 10 and 11 plot the estimated δ and α for each centile of the population-weighted distance variable. All the results discussed in section 5.1 remain qualitatively and quantitatively stable from the 50^{th} through the 75-80^{th} centile of the population-weighted distance. Most notably, the δ and the α parameters decrease and become indistinguishable from zero beyond the 80^{th} centile of distance: when classifying as *close* districts that were actually *far* from the March, it is impossible to identify any effect.

Next, I check the lack of pre-trends across the population-weighted protest buffers that convey promising DiD estimates, progressively classifying as *close* CD located within the 50^{th} through the 80^{th} centile of the population-weighted distance. Figures 12, 13 and 14 show the plot of the δ_τ (equation (4)) and α_τ (equation (5)), separately for each of the 7 population-weighted protest buffers. Results remain comparable to the estimates discussed in section 5.2 up to the 75-80^{th} centiles for the probability of having at least one female candidate and for the share of women running for office in the Republican primaries (Figures 12a and 13a). As regards results on Democratic primaries, the estimated effects remain stable for the 75^{th} and 80^{th} centiles only (see Figures 12b and 13b). Moreover, the pattern of the event-study estimates for the probability of electing a woman in the general elections closely resembles results discussed in section 5.2 (see Figure 14).

All in all, these robustness suggest that results are not driven by the arbitrary choice of the population-weighted protest buffer and should reassure the reader about the plausibility of the parallel trends across different treatment definitions.

6.2 Robustness across geographical protest buffers

Another concern related to identification could consist in the choice of 2018 population-weighted CD centroids as a proxy for the geolocation of the electorates' bulk. In this section, I assess the robustness of the results when using 2018 geographical CD centroids to proxy the geolocation of the core of the constituency. This translates to defining treatment status using geographical rather than population-weighted protest buffers. More specifically, I assess the stability of the estimated parameters by replicating the analysis discussed in section 6.1. The results of these robustness checks are in Appendix A.

I begin defining the geographical distance variable as the distance between the 2018 CD geographical centroids and the border of the nearest urban polygon hosting a March (see Table 2). Table A1 shows the distribution of the geographical distance and the corresponding geographical protest buffers: 30% of geographical CD centroids fall within protesting urban area polygons. Similarly to the population-weighted distance, the geographical distance is also overdispersed.

I then estimate equations (1) and (3) by progressively increasing the size of the treated group based on the distribution of the geographical distance. Figures A1, A2 and A3 plot of the estimated δ and α across geographical protest buffers. As regards the probability of having at least one female candidate in the parties' primaries, estimates on Republican races are comparable to the results discussed in section 5.1 across the 50^{th} and 55^{th} centiles. However, I find no evidence of an effect of the March on the probability of having at least one female candidate in Democratic races. Nonetheless, the effect of the March on the intensive margin of the supply of female politicians in Republican primaries (Figure A2) can be clearly identified discretizing the geographical distance up to the 85^{th} centile and is comparable to the results discussed in section 5.1. The same holds true for the probability of electing a female representative in general elections.

Last, I investigate the absence of pre-trends across geographical protest buffers by estimating equations (4) and (5). Figures A4, A5 and A6 plot the estimated δ_τ (equation 4) and α_τ (equation 5). Results for the probability of having at least one female candidate in Republican races show the lack of any lead effects, suggesting that $\delta_{\tau=2018}$ can have a causal interpretation across the 50^{th} and 55^{th} geographical distance centiles. Similarly, the parallel trend assumption seems to hold up to the 80^{th} centile for the share of women running in Republican races. Moreover, the parallel trends hold across the 50^{th} and 55^{th}

geographical distance centiles for the probability of electing a female US House representative in general elections.

These additional robustness suggest that results are not driven by the arbitrary choice of a geographical proxy for the geolocation of the electorate's bulk, while at the same time reinforce the ground for a causal interpretation of the estimated parameters.

7 Discussion and Conclusion

This study shows how the 2017 Women's March had a positive effect on the propensity of women to run for political offices at the federal level, ultimately reducing the gender gap in the US House. These results are consistent across flexible difference in differences parameterizations (i.e. classical DiD and event study analyses) and treatment definitions (i.e. across population-weighted and geographical protest buffers). Furthermore, the effects are heterogeneous along party lines, greater for Republican women and smaller for Democrats.

The 2017 Women's March was a massive showcase of female leadership and it favored the emergence of local resistance groups: Fisher (2019) administered surveys to protest participants throughout the wave of protests inaugurated by the March and documents how the same *resisters* that marched on January 21^{st}, 2017, engaged in local electoral campaigning for Democratic candidates. The *resisters* were mainly college educated white women, that identified themselves as Democrats and had a median age of 43 years. The peer effects of this political, and to some extent feminist, local activism are consequently identifiable on marginal conservative women, who were pushed to place candidatures in the GOP primaries *close* to the March.

The reason why conservative women were led to enter into federal politics could be that the March shaped social norms towards female leaders, reducing the social cost that women have to bear when acting as political leaders. To shed light on this mechanism, I plan to replicate the identification strategy discussed in section 4 on public opinion data from the NationScape survey (Tausanovitch and Vavreck, 2021).

A competing mechanism that could explain these findings is that the Republican Party could have favored women's candidatures in districts *close* to the March so as to adopt tactics of *strategic descriptive*

representation (Weeks et al., Forthcoming). Indeed, the March made clear that voters were demanding for more female leaders. Therefore, having a woman to list on the general election ticket could have translated to an electoral premium for the GOP. This hypothesis seems to be supported by the historical evolution of the supply of female politicians in partisan primaries (Figures 2 and 3): while we observe the *blue wave* (Fisher, 2019) that flooded Democratic primary races, there is no temporal discontinuity in the supply of female Republican candidates.

One limitation of this study is that it is difficult to disentangle women's choice to run for office from parties' choice to strategically sponsor female candidates: the informal process that leads a woman to place her candidature in the party's primary of a specific congressional district is not reflected in election records. To overcome this limit and disentangle the two components, I plan to replicate the analysis with data on female candidatures for local political offices. Indeed, the electoral process that leads a woman to enter into local politics is detached from Party politics, allowing to isolate women's choice from party's choice (Wasserman, Forthcoming).

The findings of this analysis provide evidence that grass-root feminist activism can shape the supply of female politicians in partisan primaries and have consequences on female political representation at the federal level. More research is needed to disentangle the mechanisms behind the effect identified, but the insights arising from this analysis suggest that social movements can have implications on the out-group of the organizing members.

References

Akerlof, George A, and Rachel E Kranton. 2000. "Economics and identity." *The quarterly journal of economics*, 115(3): 715–753.

America Votes 34. 2021. "America Votes 34 2019-2020, Election Returns by State."

Anastasopoulos, Lefteris. 2016. "Estimating the gender penalty in House of Representative elections using a regression discontinuity design." *Electoral Studies*, 43: 150–157.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi. 2020. "Importing political polarization? The electoral consequences of rising trade exposure." *American Economic Review*, 110(10): 3139–83.

Ballotpedia. 2020*a*. "Filing requirements for congressional candidates." Accessed 2020/09/18.

Ballotpedia. 2020*b*. "Redistricting." *https://ballotpedia.org/Redistricting*. Accessed 2020/09/18.

Bargain, Olivier, Delphine Boutin, and Hugues Champeaux. 2019. "Women's political participation and intrahousehold empowerment: Evidence from the Egyptian Arab Spring." *Journal of Development Economics*, 141: 102379.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful women: does exposure reduce bias?" *The Quarterly journal of economics*, 124(4): 1497–1540.

Broomflied, Matt. 2017. "Women's March against Donald Trump is the largest day of protests in US history, say political scientists." *Independent*. Accessed 2020/06/12.

Calderon, Alvaro, Vasiliki Fouka, and Marco Tabellini. 2021. "Racial diversity, electoral preferences, and the supply of policy: the Great Migration and civil rights." *Harvard Business School BGIE Unit Working Paper*, , (20-017).

Casas-Arce, Pablo, and Albert Saiz. 2015. "Women and power: unpopular, unwilling, or held back?" *Journal of political Economy*, 123(3): 641–669.

CAWP. 2020. "Congressional Women Candidates Database." *Center for American Women and Politics, Eagleton Institute of Politics, Rutgers University*.

Chenoweth, Erica, and Jeremy Pressman. 2017. "This is what we learned by counting the women's marches." *The Washington Post*.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1–45.

Conley, Timothy G, Christian B Hansen, and Peter E Rossi. 2012. "Plausibly exogenous." *Review of Economics and Statistics*, 94(1): 260–272.

Dave Leip's Election Atlas. 2020. "Dave Leip's Atlas of US Presidential Elections."

Davis, Angela Y. 1981. *Women, race, & class.*

Easley, Jason. 2017. "Women's March Is The Biggest Protest In US History As An Estimated 2.9 Million March." *PoliticusUsa*. Accessed 2020/06/12.

Ebbinghaus, Mathis, Nathan Bailey, and Jacob Rubel. 2021. "Defended or defunded? Local and state policy outcomes of the 2020 Black Lives Matter protests." *SocArXiv*.

El-Mallakh, Nelly, Mathilde Maurel, and Biagio Speciale. 2018. "Arab spring protests and women's labor market outcomes: Evidence from the Egyptian revolution." *Journal of Comparative Economics*, 46(2): 656–682.

FEC. 2020. "Federal Election Commission, Guide." *https://www.fec.gov/help-candidates-and-committees/guides/?tab=candidates-and-their-authorized-committees*. Accessed 2020/06/12.

Ferrara, Andreas, Patrick A Testa, Liyang Zhou, et al. 2021. "New area-and population-based geographic crosswalks for US counties and congressional districts, 1790-2020." Competitive Advantage in the Global Economy (CAGE).

Fisher, Dana R. 2019. "American Resistance: From the Women's March to the Blue Wave." *Columbia University Press*.

Fisher, Dana R, Dawn M Dow, and Rashawn Ray. 2017. "Intersectionality takes it to the streets: Mobilizing across diverse interests for the Womenâs March." *Science Advances*, 3(9): eaao1390.

Fisher, Dana R, Kenneth T Andrews, Neal Caren, Erica Chenoweth, Michael T Heaney, Tommy Leung, L Nathan Perkins, and Jeremy Pressman. 2019. "The science of contemporary street protest: New efforts in the United States." *Science advances*, 5(10): eaaw5461.

Fowler, Anthony, and Andrew B Hall. 2015. "Congressional seniority and pork: A pig fat myth?" *European Journal of Political Economy*, 40: 42–56.

Jacobson, Gary C. 2015. "It's nothing personal: The decline of the incumbency advantage in US House elections." *The Journal of Politics*, 77(3): 861–873.

Jamieson, Amber. 2016. "Women's March on Washington: a guide to the post-inaugural social justice event." *Guardian*. Accessed 2020/06/13.

Jensen, Robert, and Emily Oster. 2009. "The power of TV: Cable television and women's status in India." *The Quarterly Journal of Economics*, 124(3): 1057–1094.

Kearney, Laila. 2016. "Hawaii grandma's plea launches women's march in Washington." Accessed 2020/06/13.

Klein Teeselink, Bouke, and Georgios Melios. 2021. "Weather to Protest: The Effect of Black Lives Matter Protests on the 2020 Presidential Election." *Available at SSRN 3809877*.

Kranton, Rachel E. 2016. "Identity economics 2016: Where do social distinctions and norms come from?" *American Economic Review*, 106(5): 405–09.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap operas and fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 4(4): 1–31.

Larreboure, Magdalena, and Felipe González. 2021. "The impact of the Women's March on the US House Election." *Documento de Trabajo IE-PUC*, , (560).

Levy, Roee, and Martin Mattsson. 2021. "The effects of social movements: Evidence from# MeToo." *Available at SSRN 3496903*.

Lusher, Adam. 2016. "Donald Trump: all the sexist things he said." *Independent*. Accessed 2020/14/06.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do political protests matter? evidence from the tea party movement." *The Quarterly Journal of Economics*, 128(4): 1633–1685.

Matsa, David A, and Amalia R Miller. 2011. "Chipping away at the glass ceiling: Gender spillovers in corporate leadership." *American Economic Review*, 101(3): 635–39.

Mazumder, Soumyajit. 2018. "The persistent effect of US civil rights protests on political attitudes." *American Journal of Political Science*, 62(4): 922–935.

Mazumder, Soumyajit. 2019. "Black Lives Matter for Whitesâ Racial Prejudice: Assessing the Role of Social Movements in Shaping Racial Attitudes in the United States." *SocArXiv*.

Miller, Michael G., and Nikki Camberg. 2020. "US House Primary Election Results (2012-2018)." *Harvard Dataverse*.

Missouri Census Data Center. 2020. "Geographic Correspondence Engine."

MIT. 2020. "US House Returns 1976-2020." *MIT Election Lab*.

NCSL. 2020. "State Primary Election Systems." *https://www.ncsl.org*. Accessed 2020/10/12.

Pande, Rohini, and Deanna Ford. 2012. "Gender quotas and female leadership." *Washington, DC: World Bank*.

Pussyhat Project. n.d.. "The Pussyhat Project." *https://www.pussyhatproject.com*.

Reny, Tyler T, and Benjamin J Newman. 2021. "The Opinion-Mobilizing Effect of Social Protest against Police Violence: Evidence from the 2020 George Floyd Protests." *American Political Science Review*, 1–9.

Simplemaps. 2020. "US Cities Comprehensive Database." *https://simplemaps.com/data/us-cities*. Accessed 2020/06/15.

Sobolev, Anton, M Keith Chen, Jungseock Joo, and Zachary C Steinert-Threlkeld. 2020. "News and geolocated social media accurately measure protest size variation." *American Political Science Review*, 114(4): 1343–1351.

Swing Left. n.d.. "Swing Left." *https://swingleft.org/*. Last accessed 2022/01/22.

Tausanovitch, Chris, and Lynn Vavreck. 2021. "Democracy Fund + UCLA Nationscape Project." *October 10-17, 2019 (version 20211215)*.

UN. 2020. "Facts and figures: Women's leadership and political participation." Accessed: 2020/11/15.

US Census. 2020. "About Congressional Districts." Accessed 15/09/20.

US Census Bureau. 2020*a*. "ACS DEMOGRAPHIC AND HOUSING ESTIMATES: congressional districts."

US Census Bureau. 2020*b*. "TIGER/Line Shapefile, 2019, 2010 nation, U.S., 2010 Census Urban Area National." *https://catalog.data.gov/dataset/tiger-line-shapefile-2019-2010-nation-u-s-2010-census-urban-area-national*. Accessed 2021/05/15.

US Census Bureau. 2020*c*. "TIGER/Line Shapefile, 2019, nation, U.S., 116th Congressional District National." *https://catalog.data.gov/dataset/tiger-line-shapefile-2019-2010-nation-u-s-2010-census-urban-area-national*. Accessed 2021/05/15.

US Constitution. 2020. "Constitution of the United Stated."

US Gov. 2020. "The House Explained."

Wallace, Sophia J, Chris Zepeda-Millán, and Michael Jones-Correa. 2014. "Spatial and temporal proximity: Examining the effects of protests on political attitudes." *American Journal of Political Science*, 58(2): 433–448.

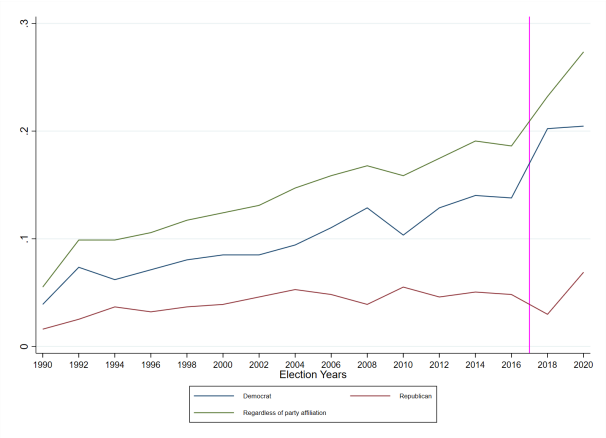
Wasow, Omar. 2020. "Agenda seeding: How 1960s black protests moved elites, public opinion and voting." *American Political Science Review*, 114(3): 638–659.

Wasserman, Melanie. 2021. "Up the Political Ladder: Gender Parity in the Effects of Electoral Defeats." *AEA Papers and Proceedings*, 111: 169–73.

Wasserman, Melanie. Forthcoming. "Gender differences in politician persistence." *Review of Economics and Statistics*.

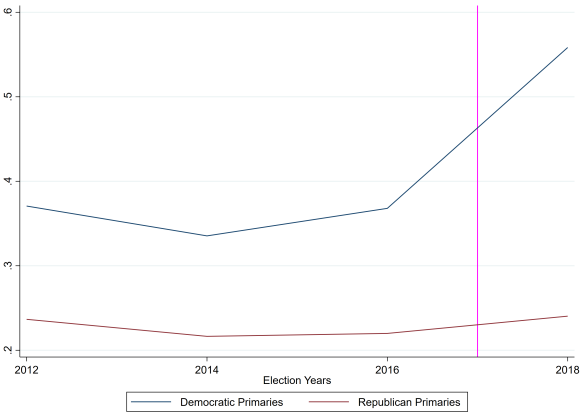
Weeks, Ana, Hilde Coffe, Bonnie Meguid, and Miki Kittilson. Forthcoming. "When Do Männerparteien Elect Women? Radical Right Populist Parties and Strategic Descriptive Representation." *American Political Science Review*.

Figure 1: Historical share of congressional districts with a female US House representative



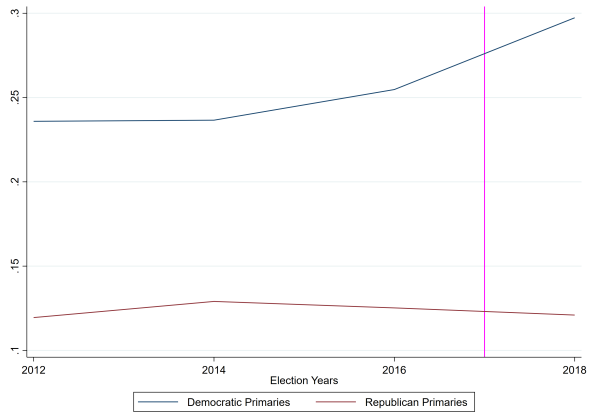
Notes: The magenta line represents the 2017 Women's March.

Figure 2: Historical share of congressional districts with at least one female candidate in primary races



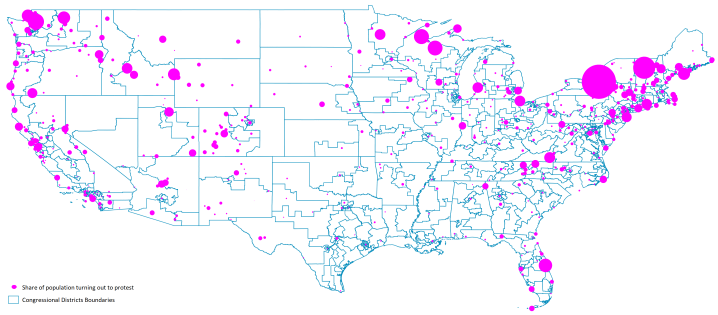
Notes: The magenta line represents the 2017 Women's March.

Figure 3: Historical share of female candidates in primary races



Notes: The magenta line represents the 2017 Women's March.

Figure 4: The historical ties of the 2017 Women's March



Notes: share of population turning out to protest in populated places. The biggest dot is Seneca Falls, the location that hosted the first American women's rights convention in 1848.

Figure 5: The 2017 Women's March, a showcase of female leadership

Source: Google Images, query *"2017 Women March"*.

Table 1: Descriptive statistics

	N	mean	sd	min	p50	p75	max
Primary Elections							
Share of female candidates:							
Republican Primaries	1361	0.124	0.273	0	0	0	1
Democratic Primaries	1373	0.257	0.363	0	0	0 .5	1
Dummy for at least one female candidate:							
Republican Primaries	1361	0.223	0.420	0	0	0	1
Democratic Primaries	1373	0.411	0.492	0	0	1	1
General Elections							
Dummy for female US House representative :							
Regardless of party affiliation	2175	0.214	0.410	0	0	0	1
Republican	2175	0.049	0.215	0	0	0	1
Democrat	2175	0.165	0.371	0	0	0	1
Protests							
Size of protests:							
Urban Protests	411	7,863.37	39,410.43	1	500	2,500	500,125
¬ Urban Protests	119	163.82	358.37	1	50	180	3000
Distance between Urban Protests and CD centroids (km):							
Population-weighted centroids	435	19.39	31.52	0	2.40	28.47	207.93
Geographical centroids	435	29.45	40.82	0	13.23	47.22	258.35
Nearest protest size:							
Size of nearest Urban Protest	435	67,047.43	124,631.3	4	7,000	52,613	500,125
Controls							
Population density (residents per km^2)	2175	899.245	2,588.876	0.4807	127.767	724.664	27,018.14
Share of votes for the Democratic party at t-1	2148	0.489	0.216	0	0.465	0.638	1

Notes: There are 495 populated places protest locations falling within urban area polygons and 119 protest locations falling outside urban areas (see maps C4 and C5). Summary statistics of the time-invariant protest variables are reported as of 2018 to allow the reader to understand that each of the 435 CD is associated with a unique distance to the nearest protest. If the CD centroid falls within the protesting urban area polygon, then the associated distance is 0 (see map C7). The table also reports the size of the nearest urban protest associated to each CD. The most densely populated CD is NY 13^{th} district in the election years 2014-16. The number of districts for which I have primary election records is smaller than the number of districts for which I have general election records. This happens because the Miller and Camberg (2020) data do not contain returns for the states that have a top-two or top-four primary system (i.e. California and Washington) and for Louisiana, where partisan primaries do not take place.

Table 2: Definition of distance variables

Variable	Definition
Population-weighted distance	Distance between the *2018 population-weighted CD centroids* and the border of the nearest protesting urban area polygon
Geographical distance	Distance between the *2018 geographical CD centroids* and the border of the nearest protesting urban area polygon

Notes: Based on the two distance variables defined in this table, I compute the corresponding distribution-based population-weighted (see Table 6) and *geographical* protest buffers (see Table A1).

Table 3: Difference in differences
Dependent variable: dummy for at least one female candidate

	Republican primaries			Democratic primaries		
	(1)	(2)	(3)	(4)	(5)	(6)
POST· *close*	0.137**	0.131**	0.119**	0.129*	0.129*	0.117*
	(0.0582)	(0.0591)	(0.0539)	(0.0702)	(0.0697)	(0.0625)
CD fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
State-election fixed effects	Y	Y		Y	Y	
State-specific linear time trends			Y			Y
Controls		Y	Y		Y	Y
Observations	1,308	1,284	1,305	1,324	1,295	1,317
R-squared	0.532	0.533	0.503	0.588	0.591	0.559
Adj R2	0.202	0.197	0.255	0.298	0.297	0.339
Dep var mean	0.224	0.224	0.224	0.357	0.357	0.357

Notes: control variables are population density and the share of votes obtained by the DP in the previous House elections. Standard errors clustered at the CD level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Difference in differences
Dependent variable: share of female candidates

	Republican primaries			Democratic primaries		
	(1)	(2)	(3)	(4)	(5)	(6)
POST · *close*	0.0738**	0.0703**	0.0633*	0.0430	0.0477	0.0584
	(0.0348)	(0.0357)	(0.0332)	(0.0483)	(0.0496)	(0.0471)
CD fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
State-election fixed effects	Y	Y		Y	Y	
State-specific linear time trends			Y			Y
Controls		Y	Y		Y	Y
Observations	1,308	1,284	1,305	1,324	1,295	1,317
R-squared	0.570	0.561	0.535	0.615	0.612	0.577
Adj R2	0.267	0.245	0.302	0.342	0.335	0.365
Dep var mean	0.122	0.122	0.122	0.241	0.241	0.241

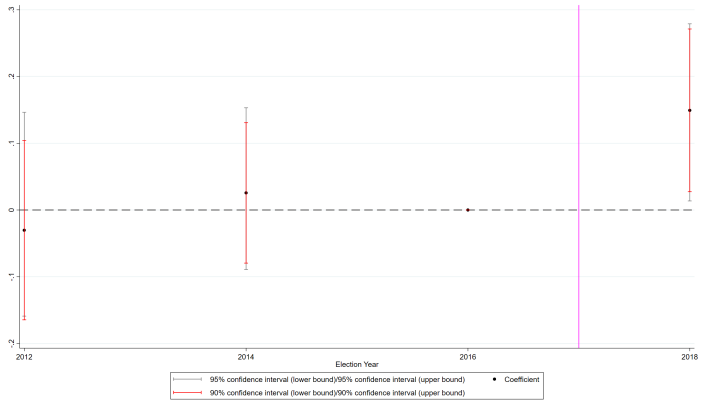
Notes: control variables are population density and the share of votes obtained by the DP in the previous House elections. Standard errors clustered at the CD level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Difference in differences
Dependent variable: dummy for female US House representative

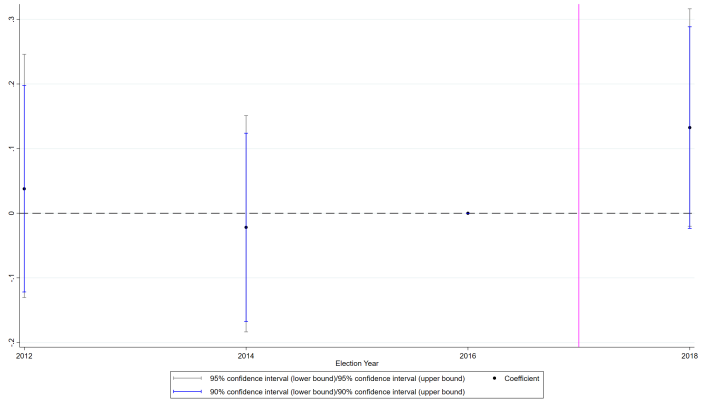
	Regardless of party affiliation			Republican			Democrat		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
POST · *close*	0.109***	0.117***	0.107***	0.0206	0.0252	0.0307	0.0888***	0.0915***	0.0759***
	(0.0372)	(0.0375)	(0.0339)	(0.0267)	(0.0273)	(0.0253)	(0.0271)	(0.0273)	(0.0234)
CD fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-election fixed effects	Y	Y		Y	Y		Y	Y	
State-specific linear time trends			Y			Y			Y
Controls		Y	Y		Y	Y		Y	Y
Observations	2,140	2,106	2,141	2,140	2,106	2,141	2,140	2,106	2,141
R-squared	0.823	0.820	0.811	0.724	0.720	0.719	0.850	0.847	0.837
Adj R2	0.736	0.731	0.746	0.589	0.582	0.623	0.776	0.771	0.782
Dep. var mean	0.185	0.185	0.185	0.0444	0.0444	0.0444	0.141	0.141	0.141

Notes: control variables are population density and the share of votes obtained by the DP in the previous House elections. Standard errors clustered at the CD level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 6: Event study
Dependent variable: dummy for at least one female candidate



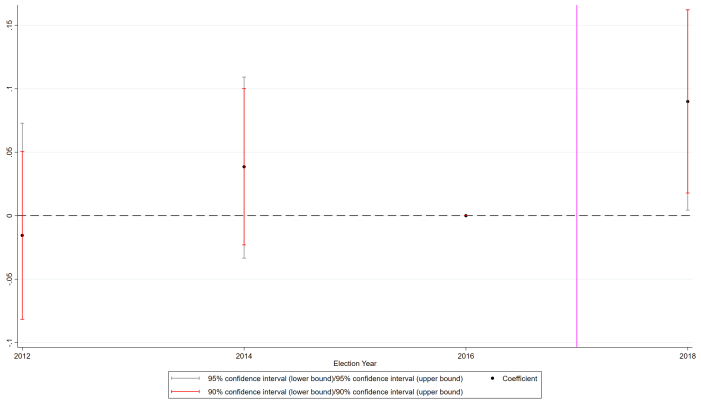
(a) Republican primaries



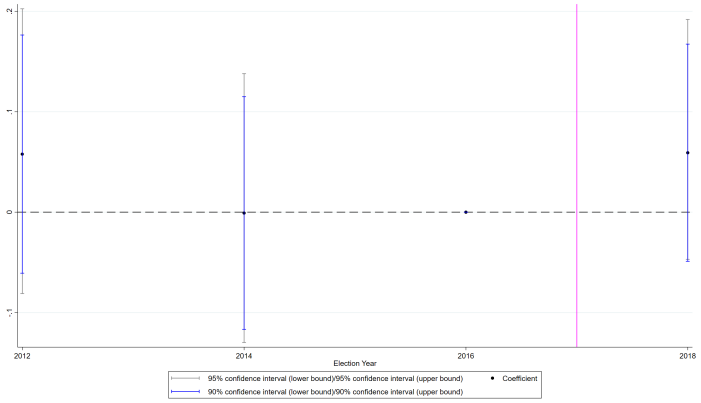
(b) Democratic primaries

Notes: plot of δ_τ, equation 4. Standard errors clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure 7: Event study
Dependent variable: share of female candidates



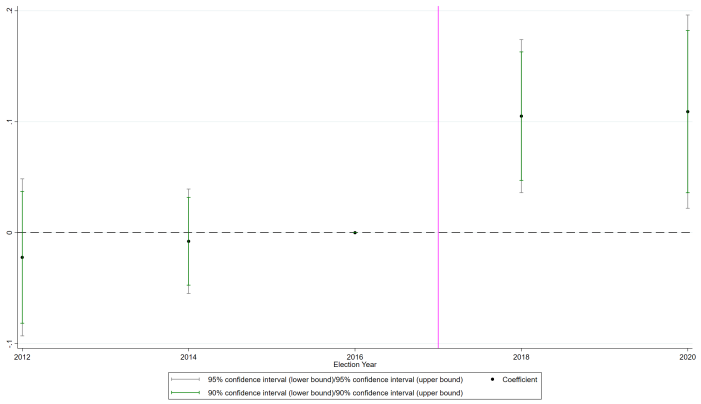
(a) Republican primaries



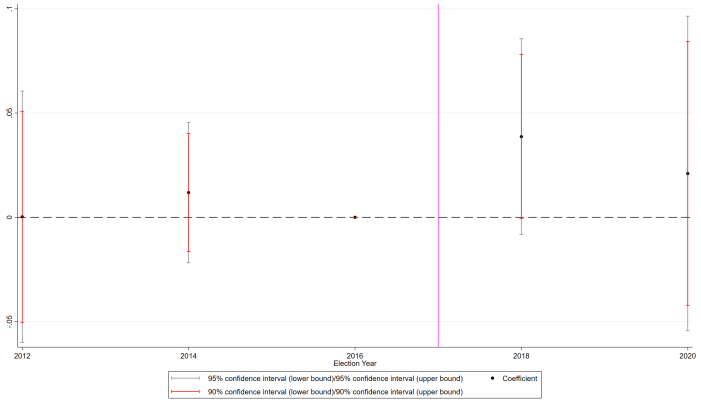
(b) Democratic primaries

Notes: plot of δ_τ, equation 4. Standard errors clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure 8: Event study
Dependent variable: dummy for female US House representative

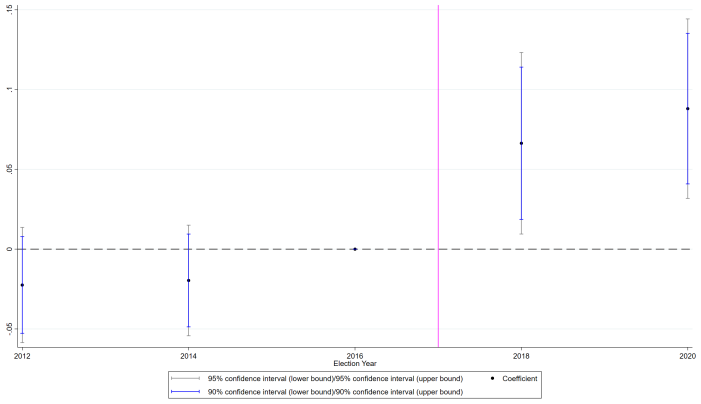


(a) Regardless of party affiliation



(b) Republican

Notes:



(c) Democrat

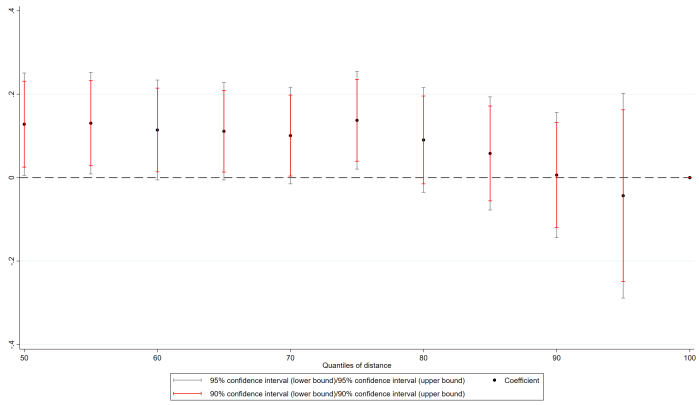
Notes: plot of α_τ, equation 5. Standard errors clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Table 6: Population-weighted protest buffers.

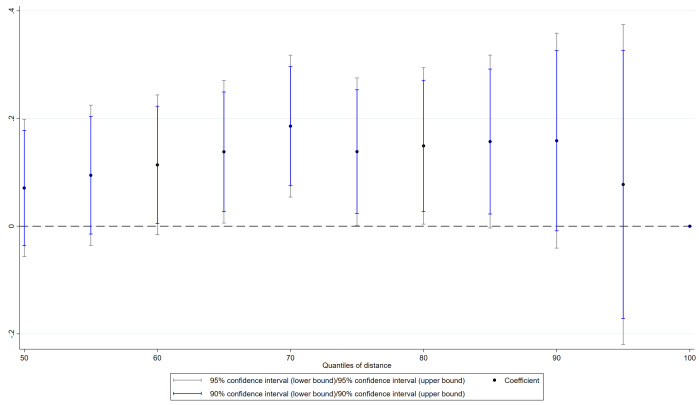
Centiles of distance	Buffers
p43	0km
p50	2.40km
p55	5.77km
p60	11.25km
p65	15.97km
p70	22.40km
p75	28.47km
p80	35.22km
p85	47.37km
p90	56.75km
p95	85.23km
p100	207.93km

Notes: The table shows the distribution of the distance between urban protests and 2018 congressional districts population-weighted centroids (i.e. the *population-weighted distance*). Each centile of the population-weighted distance corresponds to a protest buffer. 43% of 2018 CD population-weighted centroids fall within urban area polygons hosting Marches (see map C7).

Figure 9: Robustness of δ (equation 1) across population-weighted protest buffers
Dependent variable: dummy for at least one female candidate



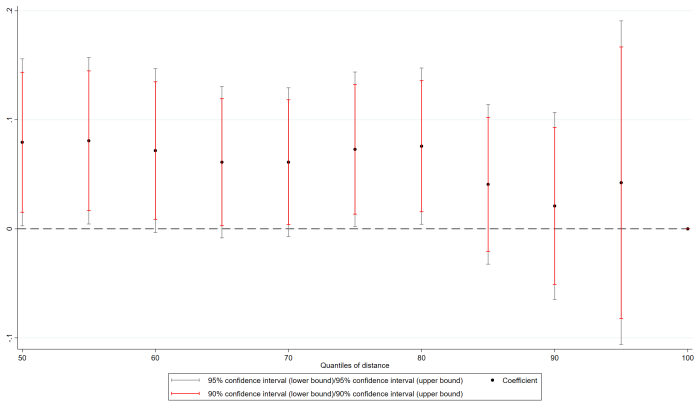
(a) Republican primaries



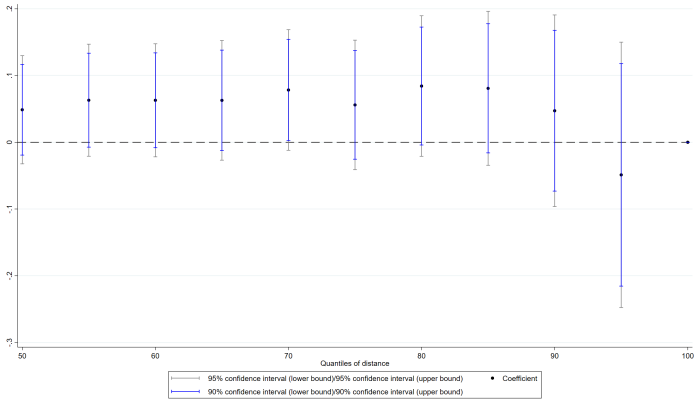
(b) Democratic primaries

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure 10: Robustness of δ (equation 1) across population-weighted protest buffers
Dependent variable: share of female candidates



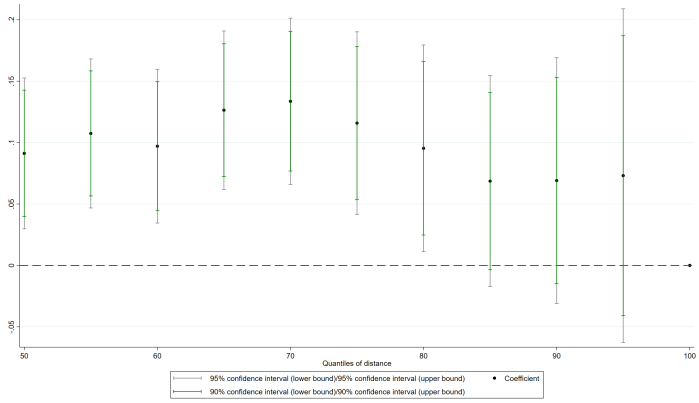
(a) Republican primaries



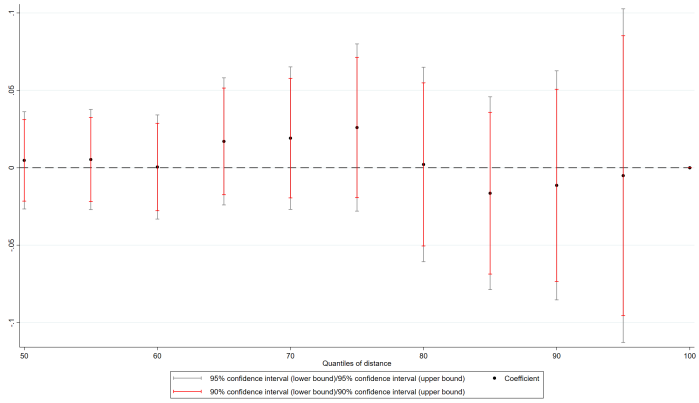
(b) Democratic primaries

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

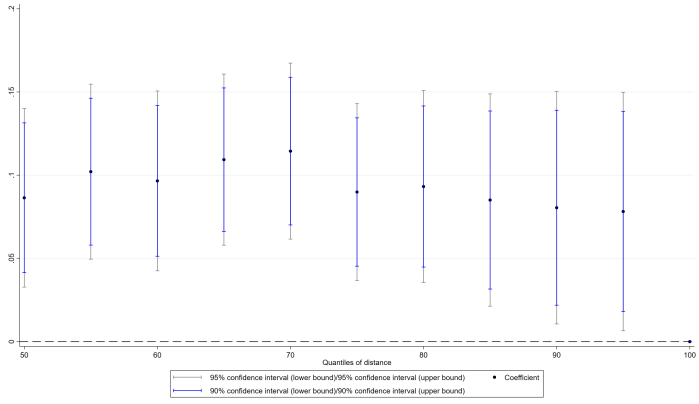
Figure 11: Robustness of α (equation 3) across population-weighted protest buffers
Dependent variable: dummy for female US House representative



(a) Regardless of party affiliation



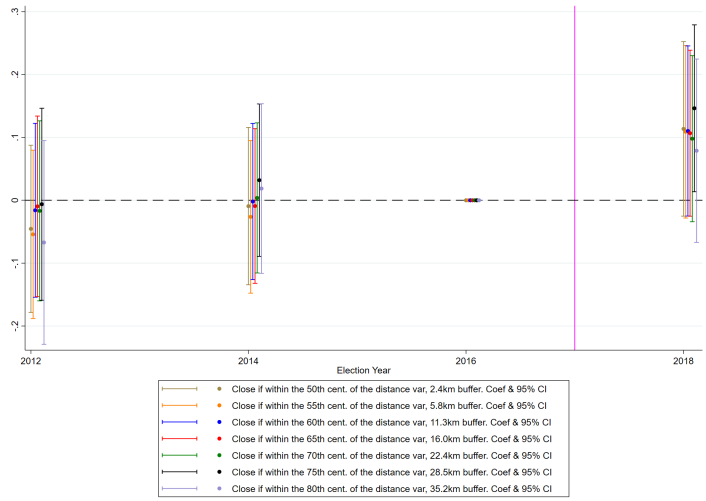
(b) Republican



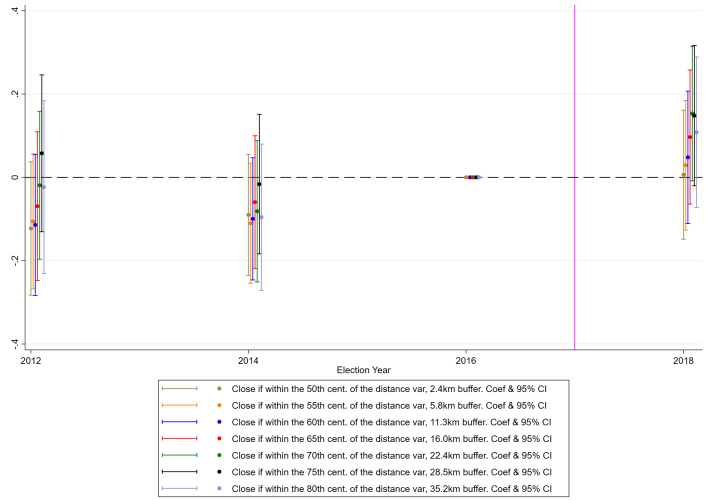
(c) Democratic

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure 12: Robustness of δ_τ (equation 4) across population-weighted protest buffers
Dependent variable: dummy for at least one female candidate



(a) Republican primaries

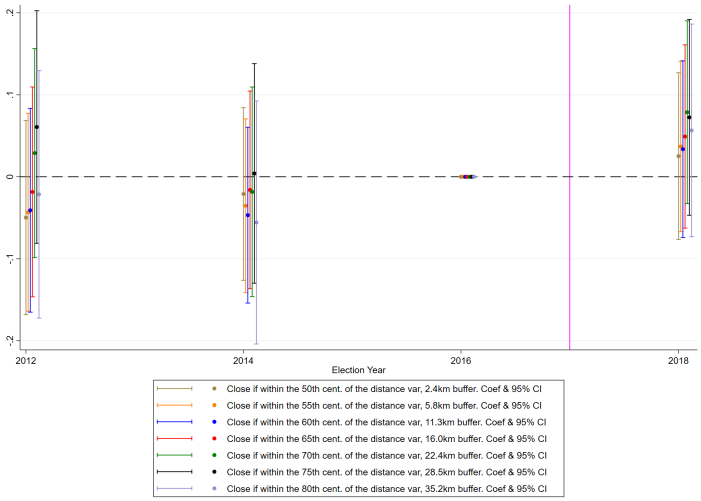


(b) Democratic primaries

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure 13: Robustness of δ_τ (equation 4) across population-weighted protest buffers
Dependent variable: share of female candidates

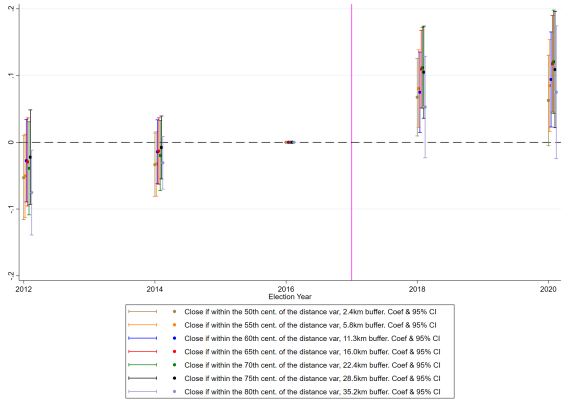
(a) Republican primaries



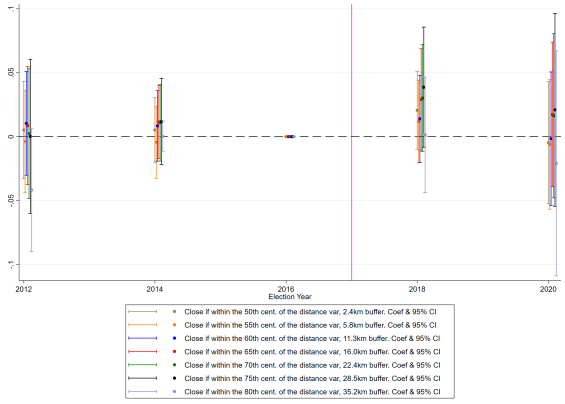
(b) Democratic primaries

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

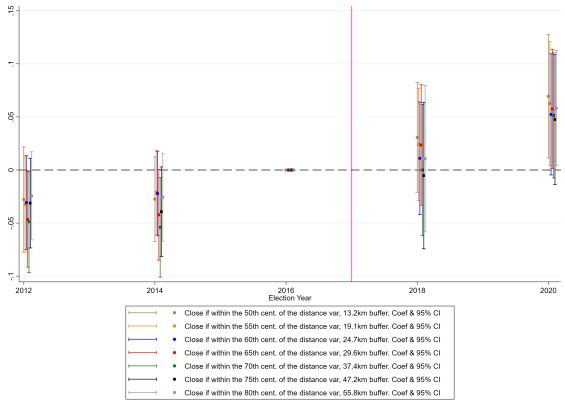
Figure 14: Robustness of α_τ (equation 5) across population-weighted protest buffers
Dependent variable: dummy for female US House representative



(a) Regardless of party affiliation



(b) Republican



(c) Democrat

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

A Appendix I - Additional robustness checks

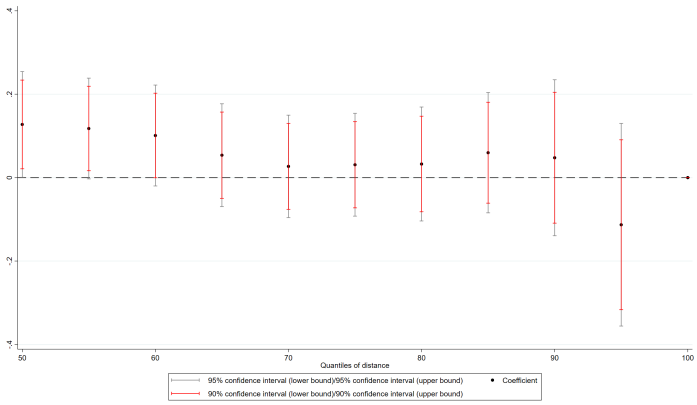
A.1 Robustness across geographical protest buffers

Table A1: Geographical protest buffers.

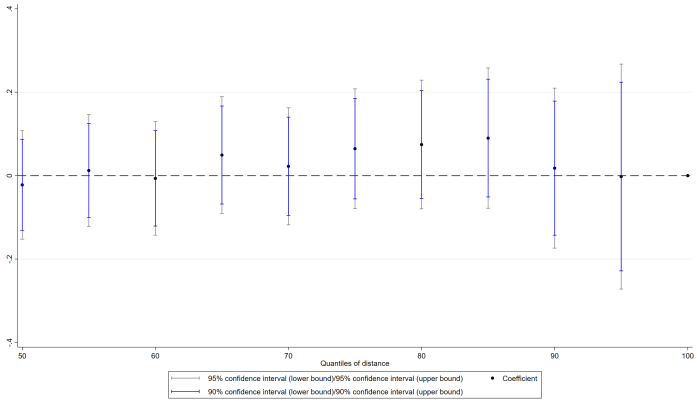
Centiles of distance	Buffers
p30	0km
p50	13.23km
p55	19.13km
p60	24.74km
p65	29.58km
p70	37.39km
p75	47.22km
p80	55.83km
p85	67.99km
p90	80.67km
p95	106.24km
p100	258.35km

Notes: The table shows the distribution of the distance between urban protests and 2018 congressional districts geographical centroids (i.e. the *geographical distance*). Each centile of the geographical distance corresponds to a protest buffer. 30% of 2018 CD geographical centroids fall within urban area polygons hosting Marches.

Figure A1: Robustness of δ (equation 1) across geographical protest buffers
Dependent variable: dummy for at least one female candidate



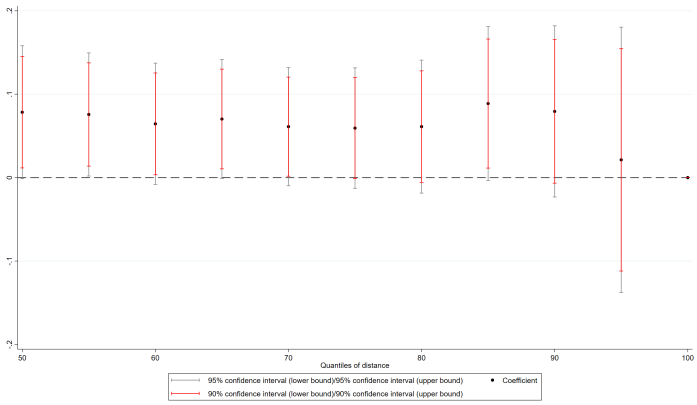
(a) Republican primaries



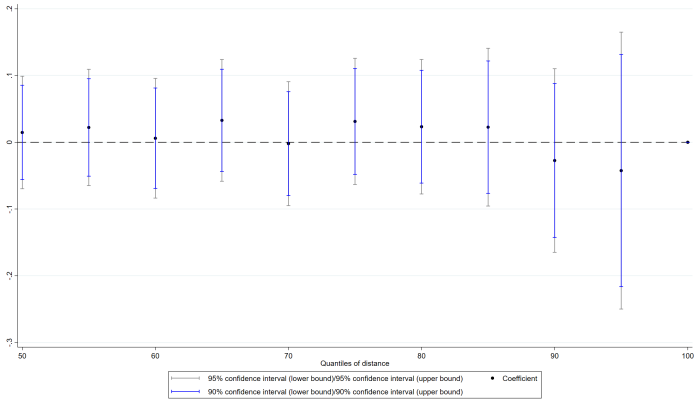
(b) Democratic primaries

Notes: districts are classified as *close* if their geographical centroid falls within the corresponding centile of the geographical distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure A2: Robustness of δ (equation 1) across geographical protest buffers
Dependent variable: share of female candidates



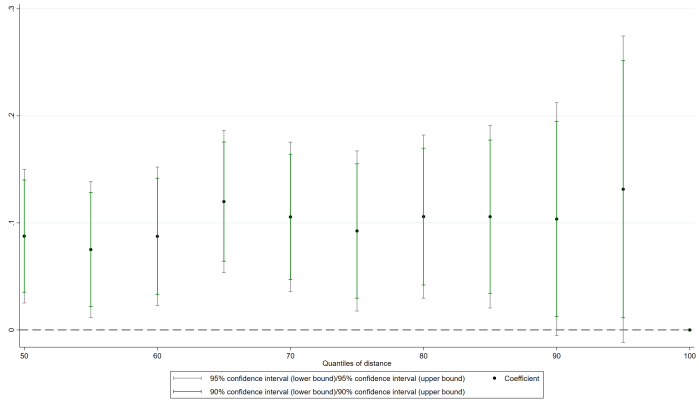
(a) Republican primaries



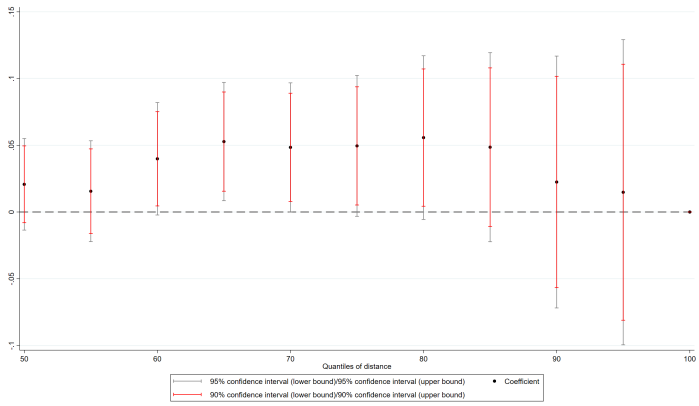
(b) Democratic primaries

Notes: districts are classified as *close* if their geographical centroid falls within the corresponding centile of the geographical distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

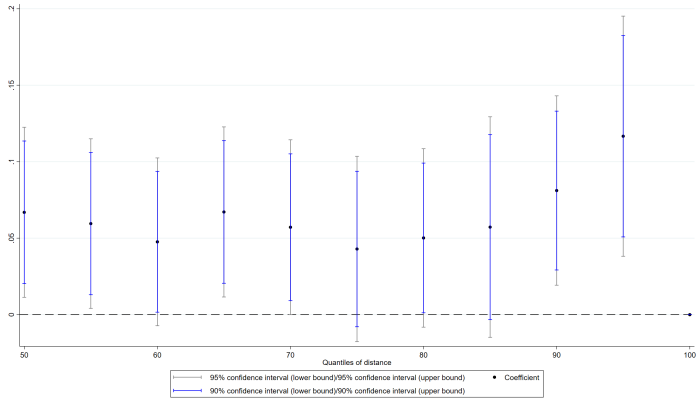
Figure A3: Robustness of α (equation 3) across geographical protest buffers
Dependent variable: dummy for female US House representative



(a) Regardless of party affiliation



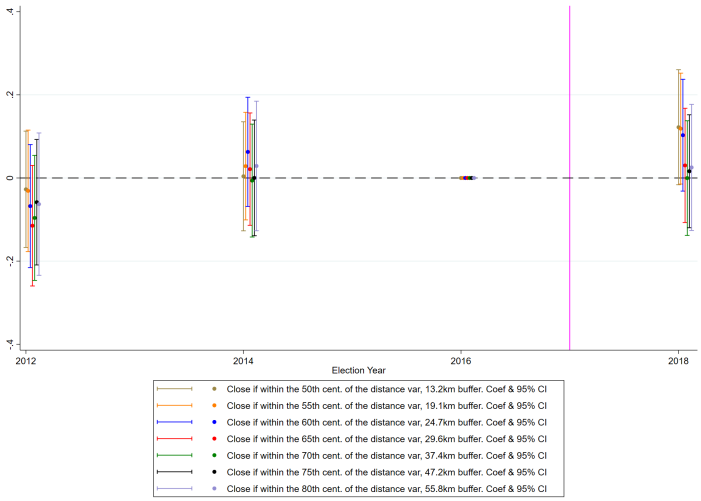
(b) Republican



(c) Democratic

Notes: districts are classified as *close* if their geographical centroid falls within the corresponding centile of the geographical distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure A4: Robustness of δ_τ (equation 4) across geographical protest buffers
Dependent variable: dummy for at least one female candidate

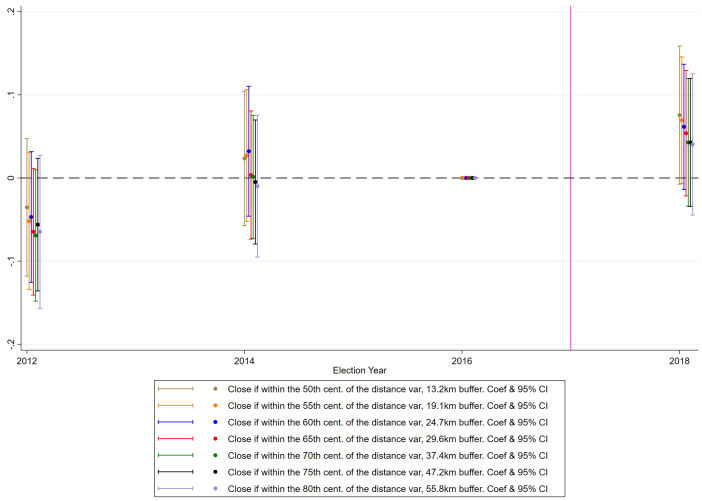


(a) Republican primaries

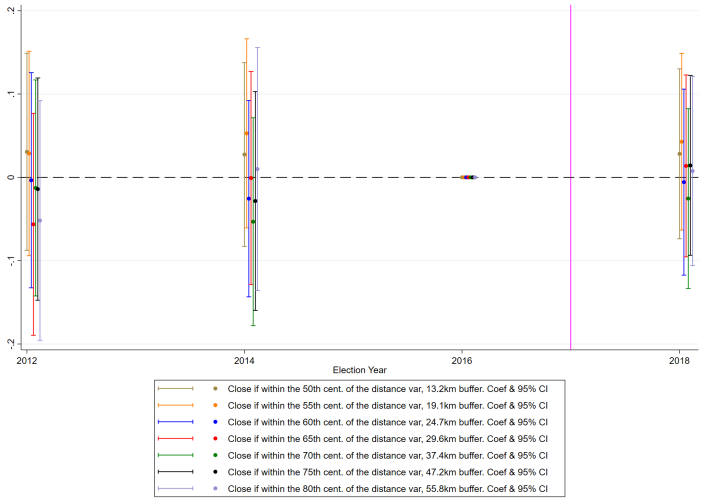
(b) Democratic primaries

Notes: districts are classified as *close* if their geographical centroid falls within the corresponding centile of the geographical distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

Figure A5: Robustness of δ_τ (equation 4) across geographical protest buffers
Dependent variable: share of female candidates



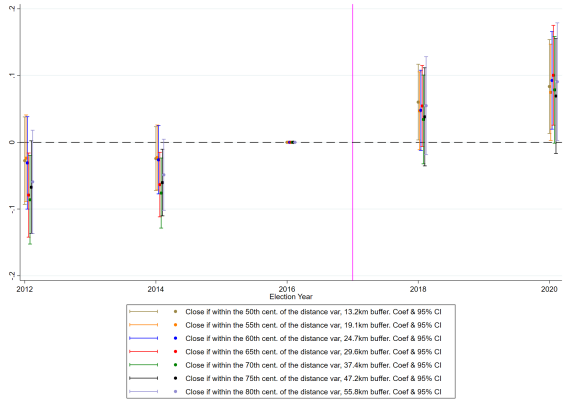
(a) Republican primaries



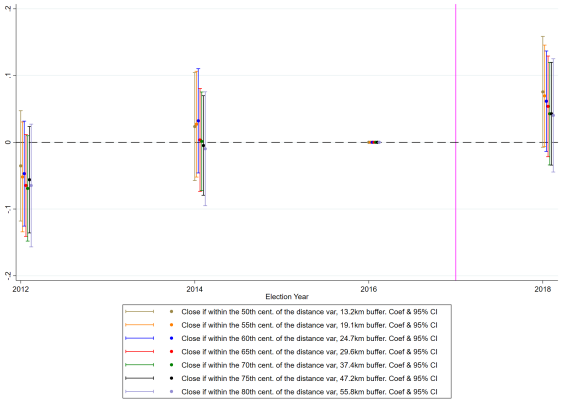
(b) Democratic primaries

Notes: districts are classified as *close* if their geographical centroid falls within the corresponding centile of the geographical distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

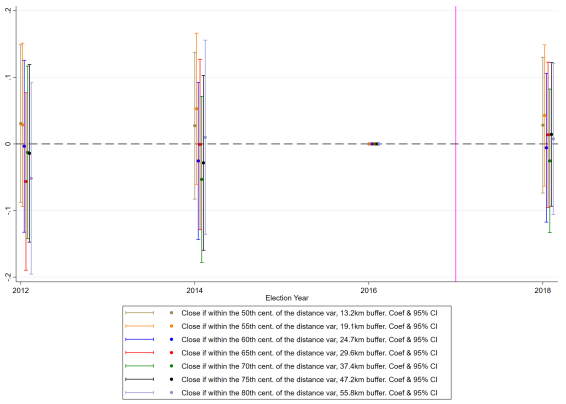
Figure A6: Robustness of α_τ (equation 5) across geographical protest buffers
Dependent variable: dummy for female US House representative



(a) Regardless of party affiliation



(b) Republican

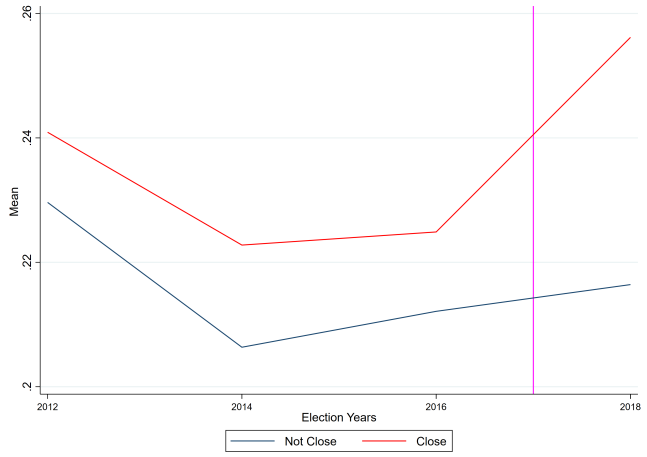


(c) Democrat

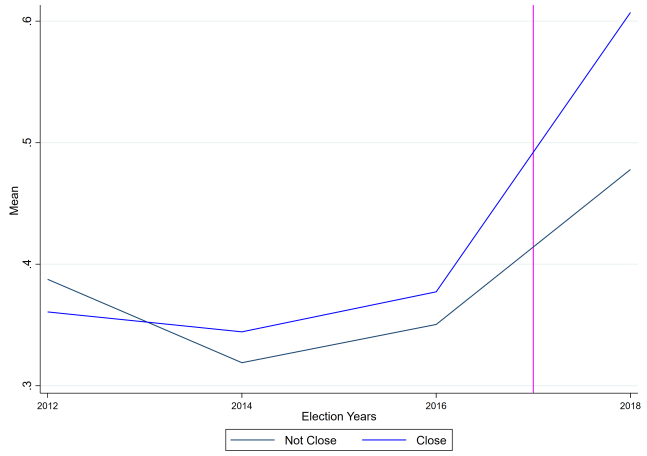
Notes: districts are classified as *close* if their geographical centroid falls within the corresponding centile of the geographical distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

B Appendix II - Descriptive figures

Figure B1: Historical share of congressional districts
with at least one female candidate in primary races by treatment status



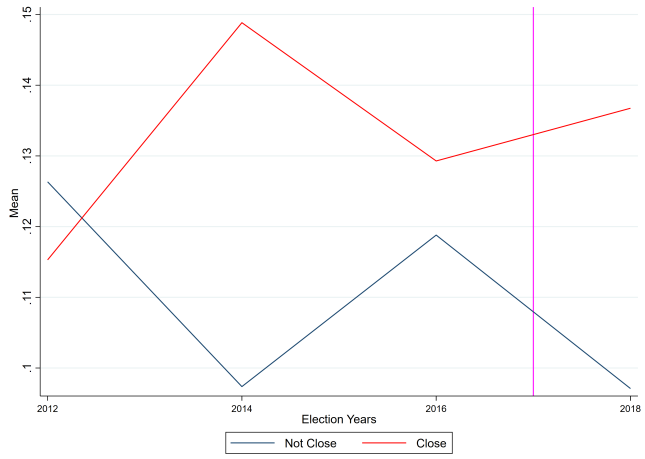
(a) Republican primaries



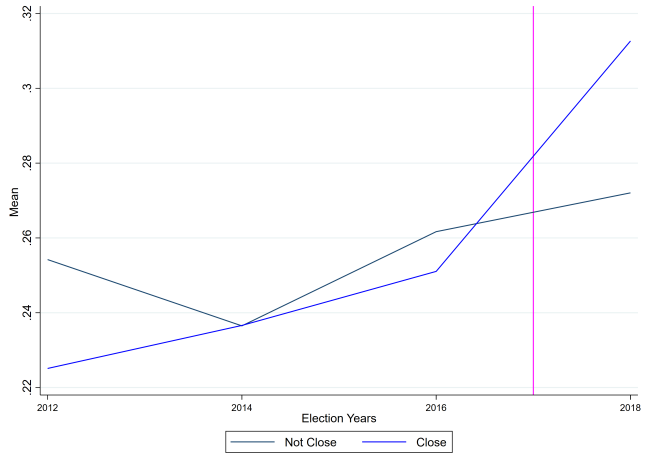
(b) Democratic primaries

Notes: The figure plots the share of CD with at least one female candidate in primary races separately for districts *close* and *not close* to the March. The magenta line represents the 2017 Women's March.

Figure B2: Historical share female candidates in primary races
by treatment status



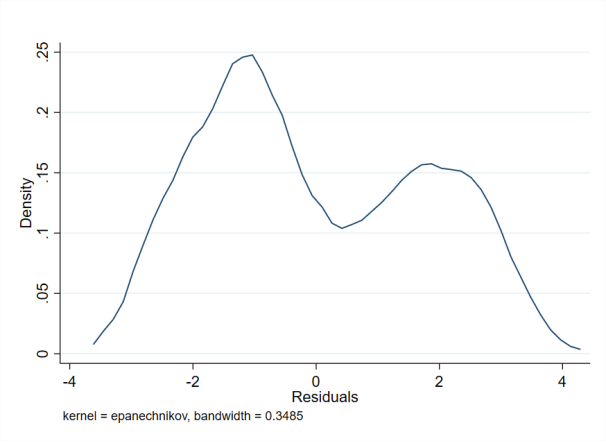
(a) Republican primaries



(b) Democratic primaries

Notes: The figure plots the share of female candidates in primary races separately for districts *close* and *not close* to the March. The magenta line represents the 2017 Women's March.

Figure B3: Bimodal distribution of the distance variable



Notes: Kernel density function of the residualized inverse hyperbolic syne of the population-weighted distance (i.e. the distance between CD's population-weighted centroids and the border of the closest protesting urban area polygon). Residuals after partialling out state fixed effects. The distribution of the residualized distance suggests that the variable can be transformed into a dummy without loosing the variability needed for identification.

C Appendix III - Other results

This section contains an ancillary analysis that replicates the empirical strategy used in this paper on general election returns.

More specifically, I analyze the impact of the March on the vote shares for women -regardless of party affiliation-, for the Democratic Party and for the Great Old Party. Such an analysis allows a direct comparison between my results and the findings of Larreboure and González (2021).

The variables analysed in this Appendix come from several data sources. The share of votes obtained by women in general elections come from the US House General Election Returns compiled by the MIT Election Lab (MIT, 2020). This data do not include information the sex of the candidate. Hence, to build the vote shares for women I follow Wasserman (Forthcoming) and use Social Security Name Files to infer candidate's sex. However, I have been able to code only 80% of the candidates in the MIT (2020) data, which translates into using the sample of candidates with most common names for making inference. The share of votes obtained by Democrats and Republicans and the corresponding dummy variables for Democrat and Republican US House representative come from the Dave Leip's Election Atlas (2020). Table C1 reports the summary statistics.

Throughout this Appendix, I report the baseline difference-in-differences and event study specifications (see section 4) and the robustness across different population-weighted protest buffers (see section 6.1.)

Table C1: Descriptive statistics of general election returns

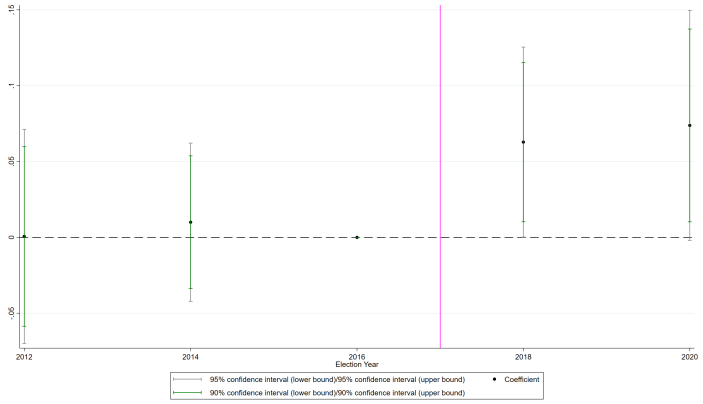
	N	mean	sd	min	max
Share of votes for:					
Women, regardless of party affiliation	2155	0.214	0.301	0	1
Republican Party	2160	0.470	0.217	0	1
Democratic Party	2160	0.497	0.214	0	1

Table C2: Difference on differences
Dependent variable: share of votes for

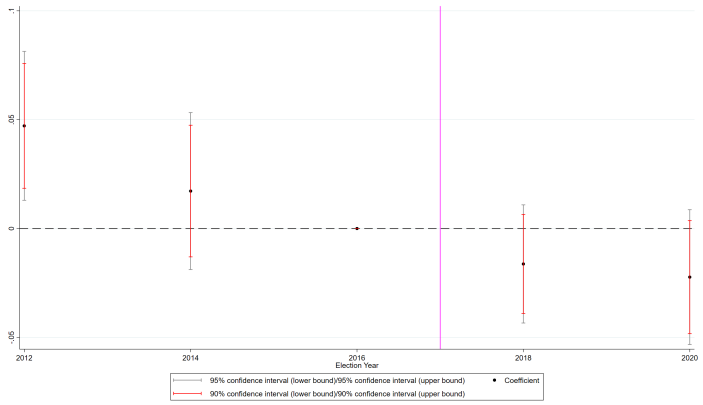
	Women, regardless of party			Republican Party			Democratic Party		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
POST · *close*	0.0654**	0.0645**	0.0622**	-0.0388***	-0.0398***	-0.0332***	0.0399***	0.0451***	0.0373***
	(0.0314)	(0.0313)	(0.0282)	(0.00991)	(0.0104)	(0.00953)	(0.0110)	(0.0118)	(0.0105)
CD fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-election fixed effects	Y	Y		Y	Y		Y	Y	
State-specific linear time trends			Y			Y			Y
Controls		Y	Y		Y	Y		Y	Y
Observations	2,119	2,086	2,121	2,118	2,092	2,127	2,118	2,092	2,127
R-squared	0.732	0.728	0.705	0.902	0.902	0.895	0.901	0.902	0.894
Adj R2	0.599	0.591	0.603	0.854	0.854	0.860	0.852	0.854	0.858
Dep. var mean	0.175	0.175	0.175	0.481	0.481	0.481	0.480	0.480	0.480

Notes: control variables are population density and the share of votes obtained by the DP in the previous House elections. Standard errors clustered at the CD level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

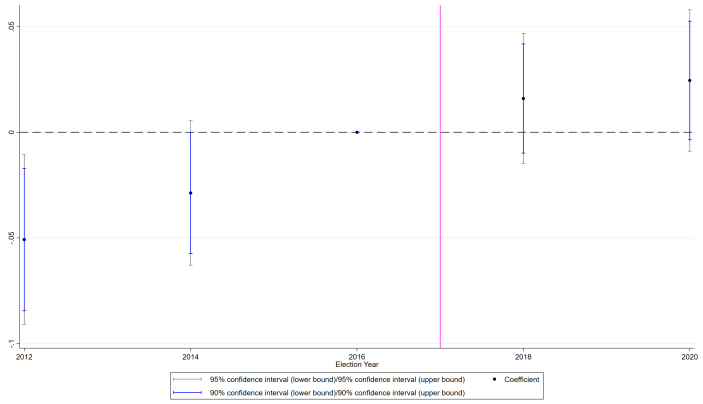
Figure C1: Event study
Dependent variable: share of votes for



(a) Women, regardless of party affiliation



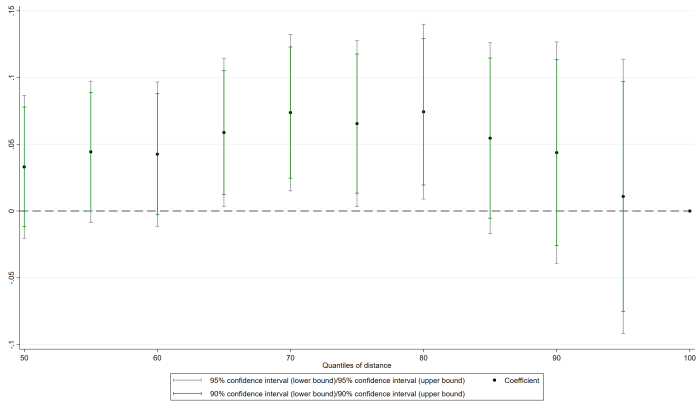
(b) Republican Party



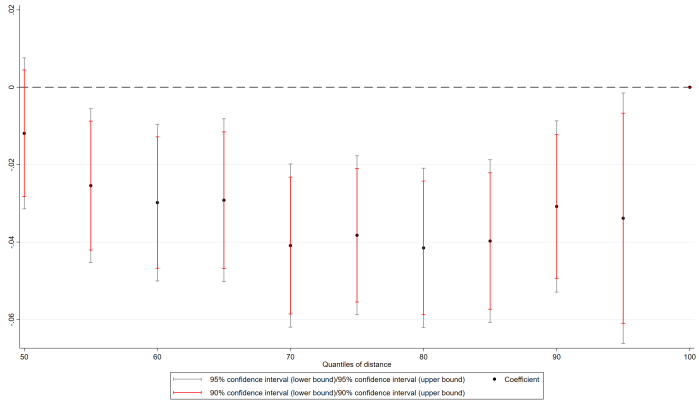
(c) Democratic Party

Notes: plot of α_τ, equation 5. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

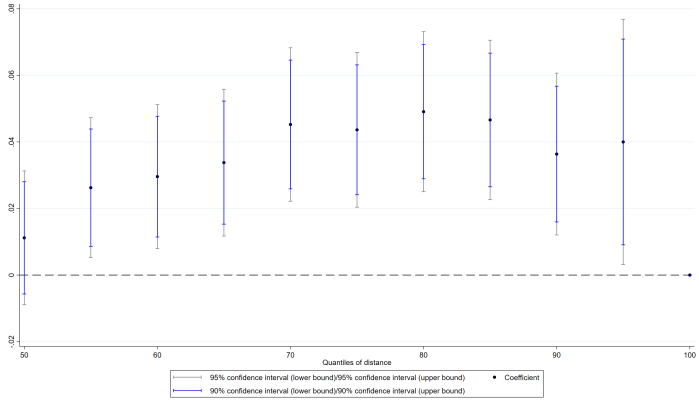
Figure C2: Robustness of α (equation 3) across population-weighted protest buffers
Dependent variable: share of votes for



(a) Women, regardless of party affiliation



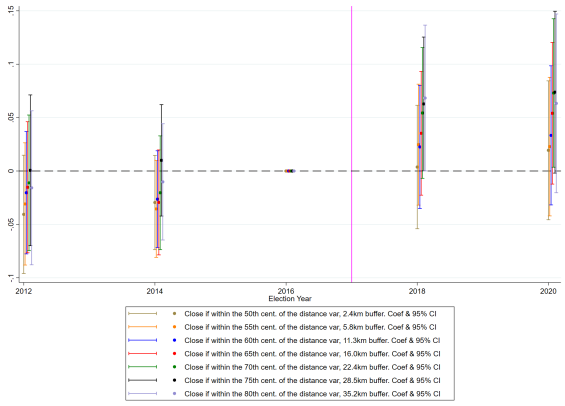
(b) Republican Party



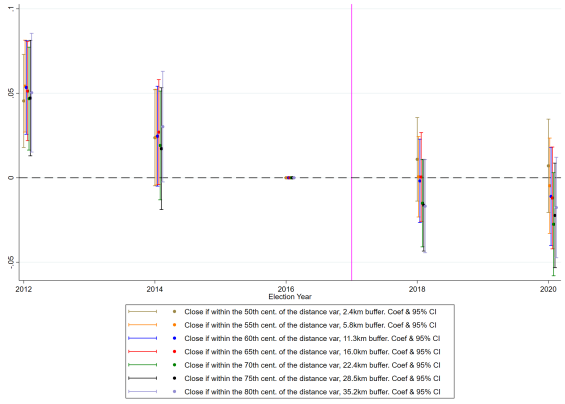
(c) Democratic Party

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

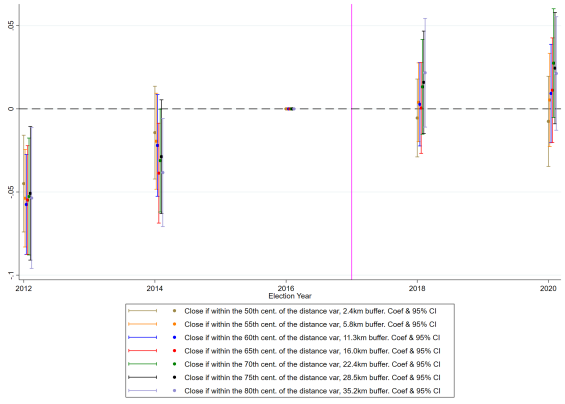
Figure C3: Robustness of α_τ (equation 5) across population-weighted protest buffers
Dependent variable: share of votes for



(a) Women, regardless of party affiliation



(b) Republican Party

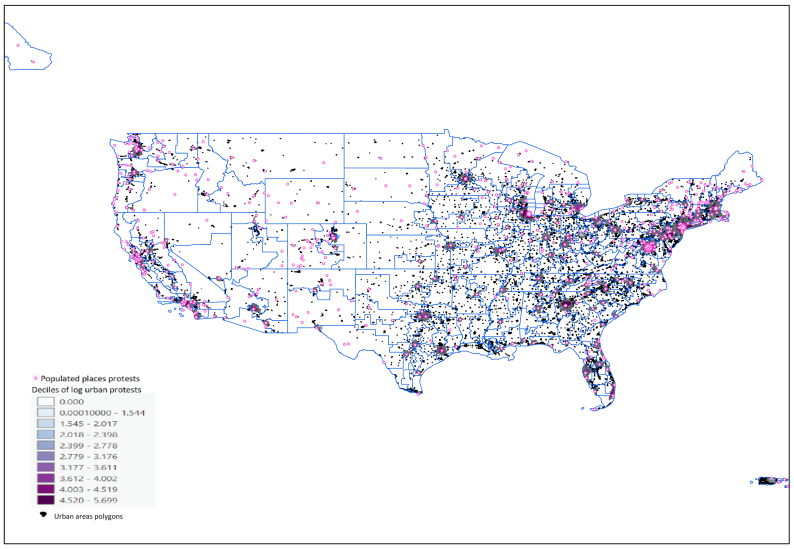


(c) Democratic Party

Notes: districts are classified as *close* if their population-weighted centroid falls within the corresponding centile of the population-weighted distance variable. Standard errors are clustered at the CD level. The reported parameters are estimated on the full specification, which includes state-election fixed effects and controls (i.e. population density and share of votes for the DP in the previous House elections).

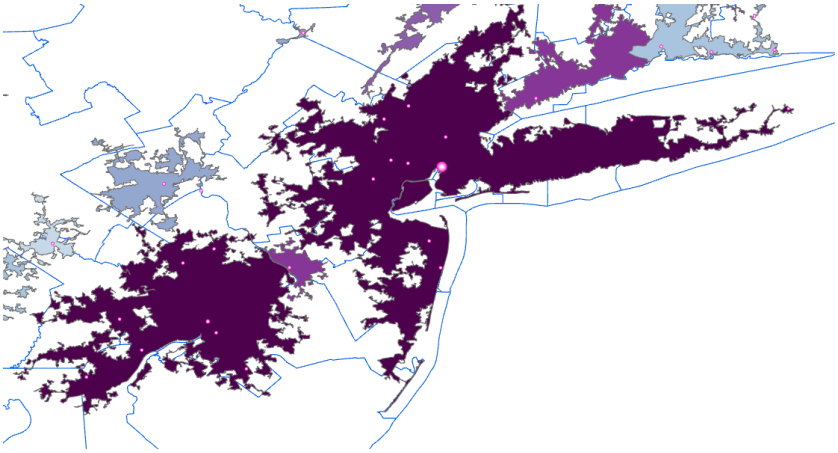
D Appendix IV - Geoprocessing

Figure C4: The geography of protests



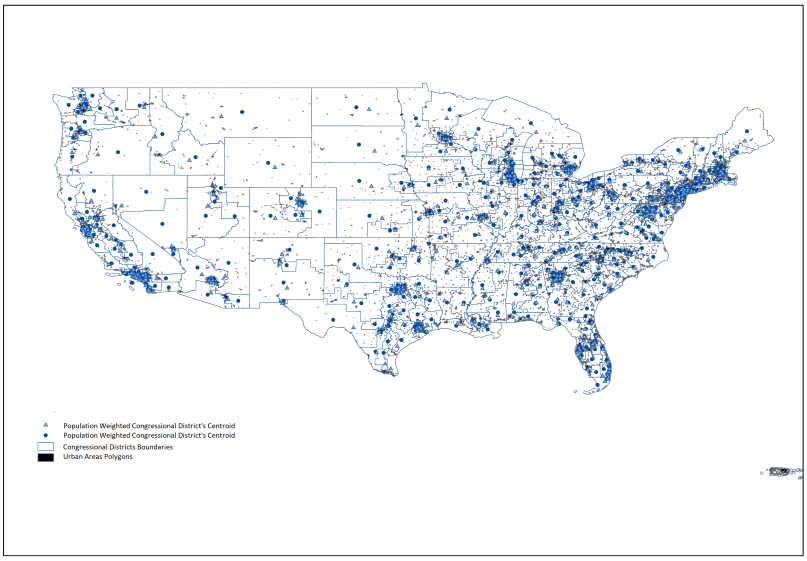
Notes: The map shows how I geographically classify protests into Urban and ¬ Urban using Tiger/Line Shapefiles. If the pink dot falls within the grey polygon, then protest is classified as Urban.

Figure C5: The geography of protests: New York



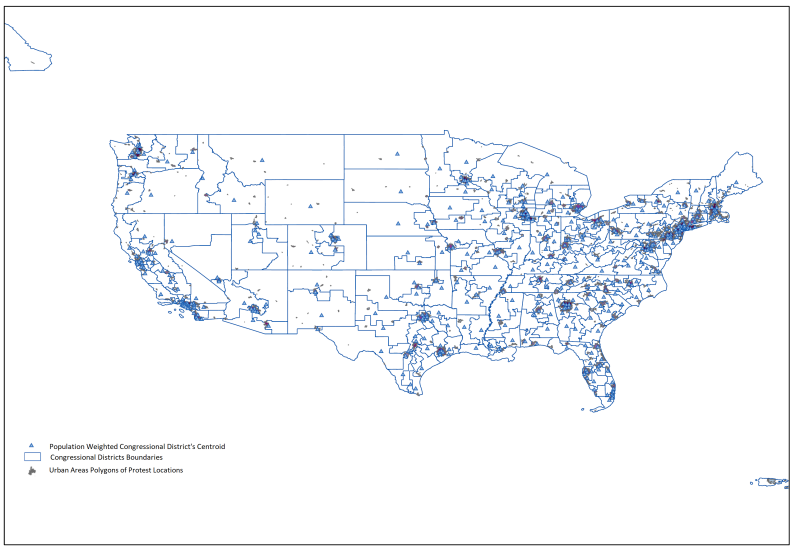
Notes: The map shows how I aggregate multiple populated places protests into a unique Urban Protest. If multiple pink dots fall within the same urban area polygon (as in the case of this map), then I take the total of protesters within the urban area polygon to infer the size of the nearest Urban Protest.

Figure C6: population-weighted and geographical 2018 congressional districts centroids



Notes: The map shows the two proxy used to assess the geolocation of the electorate's bulk. It is possible to notice that in the *"fly over zone"* population-weighted centroids (i.e. blue shaded triangles) tend to be shifted towards urban area polygons. Furthermore, in areas with greater population density (e.g. Northeastern coast) population-weighted and geographical centroids tend to overlap.

Figure C7: population-weighted centroids and Urban Protests



Notes: The map shows the overlaps between population-weighted centroids and Urban Protest. If a population-weighted centroid falls within a protesting urban area polygon, then the corresponding population-weighted distance is zero.