

# The Impact of the Women’s March on the U.S. House Election\*

*Magdalena Larrebouré*<sup>†</sup>      *Felipe González*<sup>‡</sup>

Three million people participated in the Women’s March against discrimination in 2017, the largest single-day protest in U.S. history. We ask whether the March affected the political participation of women and people from ethnic minorities in the following federal election, the 2018 House of Representatives Election. Using daily weather shocks as exogenous drivers of attendance at the March, we show that protesters were successful in empowering these historically underrepresented groups to run for a seat in Congress. Moreover, citizens of counties with more protesters increased their turnout and voted more for these groups.

*Keywords:* protests, election, gender, minority

---

\*This version: October 2019. We would like to thank Álvaro Cordero, Emilio Depetris-Chauvin, Katia Everke, Johannes Haushofer, Guillermo Marshall, Daniel Mellow, Matías Muñoz, Fernando Ochoa, Moritz Poll, Mounu Prem, José Diego Salas, Pablo Valenzuela, Cristine Von Dessauer, and Carolina Wiegand for comments and suggestions.

<sup>†</sup>Princeton University, Princeton, NJ, USA; Busara Center for Behavioral Economics, Nairobi, Kenya.

<sup>‡</sup>Pontificia Universidad Católica de Chile, Instituto de Economía, Santiago, Chile. Contact e-mail: fagonza4@uc.cl

# 1 Introduction

Three million people participated in the Women’s March of 2017, the largest single-day protest in U.S. history (Chenoweth and Pressman, 2017; Fisher et al., 2019). According to organizers, the goal was to “send a bold message to our new administration on their first day in office that women’s rights are human rights.” Previous research has shown that protests affect the policy-making process (e.g. Madestam et al. 2013), but less is known about whether these collective actions are able to empower historically underrepresented groups to run for a seat in the U.S. Congress. In this paper we estimate the local impact of protesters on the number of women and people from ethnic minorities who decided to run for office in the 2018 House of Representatives Elections. Representation matters because its impact on policies is well documented (Chattopadhyay and Duflo, 2004; Duflo, 2005; Beaman et al., 2012). Yet those who run and are elected for office rarely match the diversity in the population.<sup>1</sup> The 2018 U.S. Election was record-breaking as the number of candidates from historically underrepresented groups rose significantly (Hill, 2018).

The analysis proceeds in three steps. First, we measure the number of protesters per county using the Crowd Counting Consortium (Chenoweth and Pressman, 2017). These data aggregate information from local news, law enforcement statements, online event pages, and photos of the March. Second, we use daily weather shocks as exogenous drivers of protest attendance. Crucially, we show that after accounting for a vector of county characteristics these weather shocks are unrelated to previous political outcomes. We interpret this evidence as suggesting that the weather residuals are conditionally exogenous. Third, we use these shocks to estimate the impact of the Women’s March on the 2018 House Election. We find that protesters increased the number of candidates from historically underrepresented groups, replaced traditional candidates, and increased their vote share.

We begin by replicating Madestam et al.’s (2013) strategy, who use rainfall as an instrument for attendance to the Tea Party protest (April 15, 2009). In contrast to their findings, we show that rainfall fails to predict attendance to the Women’s March (January 21, 2017). This result can be explained by (i) differences in the geographic distribution of rainfall between the day of the Tea Party protest and the day of the Women’s March, and (ii) the different motives behind these protests. Building on their strategy and the work of Sheppard and Glassberg (2016), we create a vector with dozens of weather shocks and choose the best predictors of protest attendance using the least absolute shrinkage and se-

---

<sup>1</sup>Less than 20% of candidates were women or from ethnic minorities in the 2016 House Election, even though they represent 50 and 38% of the U.S. population (Bialik and Krogstad, 2017; Dittmar, 2018). See Dal Bó et al. (2017) for a thorough study of who becomes a politician.

lection operator proposed by Belloni et al. (2011). This “machine-chosen” weather shock is the deviation from the historical average temperature in a county-month, and it is a strong predictor of attendance to the March. Importantly, this temperature shock is uncorrelated with previous political outcomes after accounting for a vector of county characteristics.

Using the machine-chosen weather shock as an instrument for the local intensity of the Women’s March, we find that protesters increased the number of women and ethnic minorities who decided to run for Congress, contributing to the replacement of traditional candidates. More precisely, we estimate that 1,000 additional protesters, the size of the average protest in a county, are associated to one additional woman every three races and, more generally, one additional candidate from an underrepresented group every two races. This is, the number of women candidates increased by 11 percentage points or nearly 40% of the sample average. In addition, we find that underrepresented groups obtained 13 percentage points (3,000) more votes in a county, close to 32% of the sample mean, most of which is explained by an increase in the vote for white women (2,500 votes).

The main threat to our findings is a potential violation of the exclusion restriction. However, we argue that unusual weather is unlikely to have affected the election through channels different from protest attendance. A leading concern is media coverage, perhaps affected by the weather and likely to affect electoral outcomes (Strömberg, 2015). To study this possibility we checked if protests were covered by the local news in counties with the lowest and highest temperature shocks. We found local news for virtually all counties. Our findings are also robust to omit from the estimation groups of counties from the same state and outliers. Unfortunately, it is impossible to test for all possible threats. Thus we allow for a direct effect of the temperature shock and calculate that it would have to be relatively large to make the impact of protesters indistinguishable from zero (Conley et al., 2012).

This paper makes two contributions. First, we contribute to a growing literature studying historically underrepresented groups and ways to improve their representation. Our main insight is that collective actions such as protests can empower these groups to run for office and push citizens to vote for them. The majority of studies look at the case of women and estimate the impact of gender quotas, the composition of recruiting committees, and the presence of female-leadership in politics on women’s candidacies (Duflo, 2005; Beaman et al., 2009; Broockman, 2009; Bagues and Esteve-Volart, 2010; Gilardi, 2015; Baskaran and Hessami, 2018). Similarly, researchers have also studied the impact of women in politics on the selection of policies, the provision of public goods, violence against women, women’s entrepreneurship, women’s political careers, and the educational attainment of girls, finding mostly improvements in women’s lives (Chattopadhyay and Duflo, 2004; Beaman et al., 2012; Iyer et al., 2012; Ferreira and Gyourko, 2014; Ghani et al., 2014; Brollo, 2016; O’Connell,

2018, 2019). Another part of this literature focuses on similar issues but studies historically underrepresented groups different from women, both in the United States and other parts of the world (McAdam, 1982; Pande, 2003; Sass and Mehay, 2003; Banducci et al., 2004; Segura and Bowler, 2006; Preuhs, 2006; Washington, 2012; Dunning and Nilekani, 2013).

We also contribute to a literature that estimates the economic and political impacts of protests. The most recent research has shown that local collective actions such as protests and riots can affect the implementation of policies, vote shares, political attitudes, and property values (Collins and Margo, 2007; Madestam et al., 2013; Aidt and Franck, 2015).<sup>2</sup> Similarly to Madestam et al. (2013), we use geographical variation in unexpected daily weather shocks to estimate the impact of protests. In contrast to previous research, we focus on the impact of protests on the empowerment of underrepresented groups.

## 2 The Context

The Women’s March took place on January 21st 2017 and was a massive event.<sup>3</sup> Although the beginning is tied to the election of the Republican Donald Trump as President, protests were not against Republicans or Trump in particular but rather against discrimination. More than half of participants declared women’s rights to be a top motive for demonstrating, while politics was only the 8th out of 13 possible causes (Fisher et al., 2017). In between these two, protesters mentioned Equality, Reproductive Rights, Environment, Social Welfare, Racial Justice, and LGBTQIA issues. These motives point to a connection between demonstrations and a desire to improve the representation of women and other groups. In fact, according to Beyerlein et al. (2018) “[The Women’s March] reflected widely felt grievances and outrage over Trump’s election. Not only were women’s bodies being threatened, but so were the rights of immigrants, people of color, workers, and the LGBTQIA community.”

Women, African-Americans, Hispanics, Asians/Pacific Islanders, and Native Americans have been historically underrepresented in the U.S. Congress. Underrepresented groups different from women were 31% of the population but occupied only 12% of all seats in the 107th Congress in 2001. Similarly, women occupied only 13% of seats (Bialik and Krogstad,

---

<sup>2</sup>A related literature estimates the impact of *violent* protests, i.e. riots. Recent work uses modern identification strategies and finds that violence helps protesters to achieve their goals (Huet-Vaughn, 2013; Enos et al., 2019). In contrast, earlier work uses descriptive analyses and provides mixed findings (Shorter and Tilly, 1971; Welch, 1976; Snyder and Kelly, 1976; Button, 1978; Isaac and Kelly, 1981; Frey et al., 1992; McAdam and Su, 2002; Franklin, 2009; Chenoweth and Stephan, 2012).

<sup>3</sup>For example, the number of people in the Women’s March is estimated to have been over six times the number of protesters during the Tea Party Movement rallies in April 15th, 2009 (Beyerlein et al., 2018).

2017). Representation has improved but it is still far from matching the U.S. population. In this regard the 2018 Midterm Elections were record-breaking. According to studies from the Pew Research Center, the 116th U.S. Congress resulted in the most racially and ethnically diverse in American history, also breaking the record number of women serving on it (Desilver, 2018; Bialik, 2019). Overall, out of 535 members, 116 of the elected lawmakers were non-white, representing an 84% increase with respect to the 107th Congress of 2001-03. For the first time, African and Native Americans paired their share of total population with their share of Representatives in the House (12% and 1% respectively). Moreover, not only the number of congresswomen elected was the highest in U.S. history, but it was also the biggest jump in women members since the 1990s. More than a third of the 102 elected women were newcomers to the House of Representatives.

## 3 Methods

### 3.1 Data

To measure the number of protesters per county we use Erica Chenoweth and Jeremy Pressman's Data in Crowd Counting Consortium (CCC, Chenoweth and Pressman 2017; Fisher et al. 2019). The authors used publicly reported estimates of participants, validated using local news, law enforcement statements, event pages on social media, and photos of the protests. When reports were imprecise, they aimed for conservative counts. As emphasized by Fisher et al. (2019), this multisourced approach avoids problems of underreporting when using one or two newspapers (Bond et al., 1997, 2003) by allowing to check and validate the information, something particularly important for crowd counting.

The CCC reports are originally at the city level. We aggregated these to the county level and congressional district level to match the outcomes we examine. Each city belongs to a single county, hence this aggregation was straightforward. Reports were pulled together if more than one city protested within a county. In contrast, some counties lie within two or more congressional districts. To be conservative we omit these 388 counties from the analysis, ending up with 330 (out of 425) congressional districts in our sample.<sup>4</sup>

Weather data comes from the National Oceanic Atmospheric Administration (NOAA). We examine all days in January from 2011 to 2017, from nearly 6000 different weather stations in the U.S., and match each county with its nearest station. Besides a wide vector

---

<sup>4</sup>The estimates are representative of this sample and might not extrapolate to the excluded counties. Included counties are more conservative and with a higher share of white and rural population (Table A.1).

of weather variables, we follow Madestam et al. (2013) and construct variables for the amount of rain on January 21st 2017 and indicator variables for whether that day was rainy or not, using a threshold of 0.10 inches. All in all, we create a vector of 50 weather-related variables. We interpret these as weather *shocks* because we define them as the deviation from their average in January in previous years. Among these we find temperature and precipitation.<sup>5</sup> We divide temperature and rainfall shocks in bins of 2°F and 0.25 inches respectively.

We also construct demographic and electoral variables to use as controls. In terms of demographics, we follow Madestam et al. (2013) and gather county-level data for population density, income, unemployment, change in unemployment between 2013-2017, and the share of urban, Hispanic, African-American, white, and foreign-born population. Given that our focus is on the *Women’s March*, we also gather data for the share of female population, share of female citizens, and share of unmarried partners households. These data come from the U.S. Census Bureau and the American Communities Survey. We also construct log-distance from each county to Washington D.C., where the *Women’s March* took place, and electoral variables. For the latter we use the 2016 U.S. Presidential Election and 2014 House of Representatives Election. The variables comprehend Trump’s and Clinton’s vote shares, the Republican and Democratic Party vote shares and turnout per county population.

The outcomes are related to the 2018 House of Representatives Elections, data we gather from the Harvard Dataverse (Pettigrew, 2018). We observe the names of all candidates, their political parties, and turnout. We construct six outcome variables. The first three outcomes are measured by district: (i) the total number of candidates, (ii) the share of candidates who are women, and (iii) the share of candidates who belong to an underrepresented group. The remaining three outcomes are measured by county: (iv) the vote share obtained by women, (v) the vote share obtained by candidates from underrepresented groups, and (vi) turnout. The underrepresented groups in this study include women, African-Americans, Hispanics, Asian/Pacific Islanders, and Native Americans. We use districts and counties in separate samples because, although we observe protests at the city level (smaller administrative unit), the first two outcomes were decisions made at the district level, and the latter three outcomes were county-level decisions.

Table 1 presents summary statistics for counties with protesters during the *Women’s March* and counties with zero protesters. Counties with protests have a lower share of family households, less rural and white population, and a larger share of foreign born and Hispanic population. These counties are also more educated and with higher median income and less unemployment. Politically, counties with and without protests have similar turnout,

---

<sup>5</sup>We use average and maximum temperature and exclude minimum temperatures because they usually occur during the night and protests take place during the day.

but the former are more Democrat.<sup>6</sup> Therefore a simple comparison of counties with and without protests is unlikely to reveal the political impact of the Women’s March.

### 3.2 Empirical strategy

To estimate the impact of the Women’s March we use an instrumental variables framework. The relationship of interest can be written as follows:

$$Y_i = \alpha + \beta \cdot \text{Protesters}_i + x'_i \delta + \epsilon_i \quad (1)$$

where  $Y_i$  is an outcome of interest in county  $i$ ,  $\text{Protesters}_i$  is a measure of protest intensity,  $x_i$  is a vector of predetermined control variables, and  $\epsilon_i$  is a mean-zero error term. As discussed, a naive OLS estimation of  $\beta$  is unlikely to represent the causal effect of protests because of omitted variables and measurement error in the number of protesters. An instrumental variables strategy can help to overcome both concerns.

Unusual weather the day of the Women’s March is likely to have an impact on protest attendance and, we argue, it is also likely to be uncorrelated with other factors driving attendance to the Women’s March and electoral outcomes. The former condition is testable, but the latter is ultimately an (identification) assumption. As argued by Madestam et al. (2013), there are two leading concerns regarding this assumption. First, weather shocks are likely to affect press coverage of the protest. Second, the weather might affect protesters’ experience during the event and affect the spread of the movement. The next section discusses why both of these threats are unlikely to be relevant in this context.

To begin the analysis we replicate Madestam et al. (2013)’s first stage strategy:

$$\text{Protesters}_i = \phi + \beta \cdot \text{Rain}_i + \zeta \cdot \text{Likelihood of Rain}_i + x'_i \lambda + \varepsilon_i \quad (2)$$

where  $\text{Protesters}_i$  is a measure of attendance to the march in county (or district)  $i$ .  $\text{Rain}_i$  is an indicator if there was at least 0.1 inches of rain the day of the event, or the amount of inches of rain fallen that day.  $\text{Likelihood of Rain}_i$  is a flexible control for the probability of rain calculated using daily weather data from previous years. The vector  $x_i$  contains pre-determined county characteristics, including past elections outcomes and demographic characteristics. Estimates are weighted by population when the protesters variable is measured per population. Standard errors are clustered at the state level, but results are robust

---

<sup>6</sup>Table A.2 presents additional characteristics which shows counties with protests are also more urban and more unmarried partners live there.

to adjusting standard errors for spatial correlation with a distance cutoff of 100 kilometers. Since rainfall is likely to decrease attendance to the rallies, we expect  $\hat{\beta}$  to be negative.

The effect of rainfall on protest attendance depends on the geographic distribution of rain that day. A more robust strategy is to follow Sheppard and Glassberg (2016) and use weather shocks selected by a data-driven algorithm. We use the least absolute shrinkage and selection operator (LASSO) method proposed by Belloni et al. (2011) to select weather instruments from a set of 50 weather shocks. In particular, we estimate:

$$\text{Protesters}_i = \omega + \beta \cdot \text{Weather Shock}_i + w_i' \lambda + \varepsilon_i \quad (3)$$

where  $\text{Weather Shock}_i$  are the LASSO-chosen instruments. The chosen variable is the standardized temperature shock the day of the march, in both the county- and district-level specifications.<sup>7</sup> Figure 1 presents a map with the variation of this shock after removing the variation from the vector of machine-chosen control variables.<sup>8</sup>

Importantly, the machine-chosen weather shock has little empirical relationship with previous electoral variables. Table 1 presents estimates of equation (3) using county characteristics as dependent variable. To avoid cherry-picking  $w_i$  these are also LASSO-chosen. The estimates reveals that  $w_i$  is important because (i) most electoral differences across counties disappear after including  $w_i$  in the estimation (column 6), and (ii) the weather shock affected counties with less foreign population, more African Americans, and more Hispanics. Therefore all specifications will include machine-chosen controls for each dependent variable.<sup>9</sup>

---

<sup>7</sup>In particular, this shock is defined as  $z_i \equiv \frac{x_i - \bar{x}_i}{\sigma_i}$ , where  $x_i$  is the average temperature in county (or district)  $i$  the day of the Women's March and  $\bar{x}_i, \sigma_i$  are the average and standard deviation of  $x_i$  calculated using five random days in January during the seven years before the March. Table A.3 presents the vector with all possible weather shocks to be chosen.

<sup>8</sup>Figure A.1 shows the geographic distribution of the temperature shock without residualizing. This map reveals spatial correlation in the temperature shock. To address this potential threat to inference in the appendix we show that results are robust when excluding one state at the time, when we cluster standard errors by state in all specifications, and when we allow errors to be correlated spatially with different geographic cutoffs using Conley's (1999) method.

<sup>9</sup>There are 24 socio-economic and 10 electoral predetermined variables to be potentially chosen as controls. Table A.4 presents all of these and Table A.5 shows the set chosen for each outcome. Results are similar if we use the controls employed by Madestam et al. (2013).

## 4 Results

### 4.1 Weather shocks and attendance to the Women’s March

Table 2 presents estimates of equations (2) and (3). Columns 1-4 replicate Madestam et al.’s (2013) econometric strategy using the number of protesters in the county over population as the endogenous variable. Rainfall the day of the event has little predictive power on the size of local protests and, if anything, the sign of the relationship is the opposite of what we expected. We highlight two possible explanations for this (null) result. First, the randomness of daily weather shocks means that the set of counties affected by it might be different during the inaugural protest of the Tea Party Movement and the Women’s March. The upper maps in Figure 1 show a different geographic distribution of shocks during these dates. Second, the Women’s March was six times larger than the Tea Party protest. Hence, the sensitivity of attendance to rainfall might differ due to the differential motives behind each protest, their size, and the time of the year in which they took place.

In contrast to the rainfall shock, the machine-chosen weather shock has a strong predictive power on protest participation (see Table 2 columns 5-6). Results indicate that a one standard deviation ( $\sigma$ ) increase in the temperature shock (0.84) decreases the share of protesters in the population by 0.43 percentage points (pp.) and 0.60 pp. in the county- and district-level regressions respectively. These coefficients represent a 41% and 57% change with respect to the sample average. The corresponding  $F$ -statistics are 18 and 34.<sup>10</sup> Why is protest attendance lower with unusual larger temperatures? Our interpretation is that with warmer temperatures in the winter the relative price of protesting increases when compared to the price of spending time with friends or family. A larger temperature shock might make protesting less attractive because of an increase of the opportunity cost.

### 4.2 The impact of the Women’s March

Table 3-A shows the direct effect of the machine-chosen instrument on the six outcomes of interest, three related to the presence of candidates in the ballot, two related to vote shares, and turnout. Panel B uses the instrument to estimate the impact of protesters, and panel C

---

<sup>10</sup>Table A.6 show that these results are similar if we measure the number of protesters in thousands or in logarithms. Figure A.2 shows that the non-parametric relationship between the temperature shock and protest attendance is approximately linear across the shock distribution.

shows OLS results for comparison.<sup>11</sup> We begin by noting the lack of an empirical association between weather shocks, protesters, and the number of candidates in the ballot (column 1). Panel A indicates that a one standard deviation increase in the temperature shock on January 21st (0.84) decreased the share of women running for Congress by 6 pp. and the share of underrepresented minorities by 8 pp. Moreover, women's vote share decreased by 3 pp., the vote share of underrepresented groups decreased by 5 pp., and turnout decreased by 0.8 pp. In terms of magnitude, each of these estimates represent changes of 24%, 21%, 12%, 10% and 2% of the sample means respectively.

Two-stage least squares estimates in panel B indicate that the Women's March had an impact on (i) the number of candidates from underrepresented groups, and (ii) their electoral outcomes. As the number of candidates remained similar, we conclude that underrepresented groups *replaced* traditional candidates. To gauge the magnitude of these estimates let us consider an increase of 1 pp. in the share of protesters in a county, approximately 1,000 more protesters and the size of the average protest. The estimates suggest that these additional protesters increased the percentage of women candidates by 11 pp. and the percentage of candidates from underrepresented groups by 17 percentage points. In terms of the ballot, these numbers mean that there was one additional women candidate every three ballots and one additional candidate from underrepresented groups every two ballots. Besides motivating people to run for Congress, protesters also increased their vote share by 13 pp. (3,000 votes) in the average county. Most of this increase is explained by an increase in the vote share of women, who got 10 percentage points (2,500) more votes. Finally, column 5 shows that protests also motivated citizens to vote, increasing turnout by 1.6 pp. or 1,600 votes.<sup>12</sup>

To understand who benefited the most from the Women's March, Table 4 splits electoral results by racial and ethnic groups. The estimates reveal that the impact of the March was mostly driven by increases in the votes for Non-Hispanic and Non-African-American Women. Put differently, in places with more protesters the additional political support for underrepresented groups favored mostly white women, African-American men, and Hispanic men. Part of these results can be explained by an increase in the number of candidates from these groups, suggesting that their additional presence in the ballot was critical.

---

<sup>11</sup>Table A.8 presents the same results using Madestam et al.'s (2013) controls plus a vector of women-related variables and estimated coefficients are virtually the same.

<sup>12</sup>Table 3-C reveals that a naïve OLS estimation delivers an attenuated coefficient. This could be explained by classical measurement error in the number of protesters or by omitted variables, e.g. protests were presumably larger in places where discrimination is harder to change in the non-protesting population.

### 4.3 Alternative explanations

Our analysis assumes that unusual weather on January 21st of 2017 affected the 2018 election only through attendance to the Women’s March. A leading concern relates to the media: unusual weather can also change media coverage, which in turn affects electoral outcomes (Snyder and Strömberg, 2010; Strömberg, 2015). This concern is unlikely to threaten our results for two reasons. First, media coverage of the March should be less affected by unusual weather than other protests in the past because of its contextual relevance and the rise of the internet. In this sense, the fact that rainfall has little impact on attendance is reassuring of the March’s importance. Second, we investigated media coverage of the March in places with weather shocks above the 90th percentile and below the 10th percentile and found media reports for all protests but two.<sup>13</sup>

Another concern relates to how temperatures affect the social experience of protesters at the protest. A large literature has shown that unusually high temperatures make humans more violent (Hsiang et al., 2013). Violence could affect the protesting experience or affect its effectiveness. This is unlikely to be a concern in our case because we find that people are *less* likely to join the Women’s March with high temperatures. In line with this statement is the fact that 95-99% of all protests were peaceful and arrest-free (Fisher et al., 2019).

Unfortunately, we cannot prove if unusual weather affected elections *only through* attendance to the March. Thus we also calculated the change in our estimates if the instrument had a *direct* impact on electoral outcomes (Conley et al., 2012). To make the impact of the March non-different from zero, the direct effect of the instrument would have to be 60, 57, 18, 52, and 48% of the reduced form effects for the five main outcomes (excluding total candidates). Because these direct effects are non-negligible, we conclude that our estimates of the March’s impact are robust to small deviations from the identification assumption.<sup>14</sup>

Lastly, we show that results are robust to the exclusion of groups of counties.<sup>15</sup> First, the March’s impact is virtually the same in 50 complementary estimations where each time we exclude all counties from a state. The exception is perhaps the case of California where estimates become larger. California experienced a low temperature shock, high attendance to the March, and it is highly populated, all of which contribute to this effect. Second, the

---

<sup>13</sup>Tables A.9 and A.10 present results. This evidence is only suggestive because it reflects the extensive margin, i.e. counties with versus without local coverage. However, the total number of media outlets covering the protests could also be important but we unfortunately lack that data to check for this empirically.

<sup>14</sup>Figure A.3 provides more details about this exercise and the full set of results.

<sup>15</sup>Figure A.4 presents the robustness of two-stage least squares estimates and, for completion, Figure A.5 presents the first-stage at the district and county levels. Table A.11 shows estimates omitting outliers.

impacts of the March are also robust to the exclusion of outlier counties. To implement this exercise we omit from the estimating sample all counties for which  $|DFBETA_i| < \frac{2}{\sqrt{N}}$ , where  $N$  is the number of observations and the term in absolute value represents the difference between estimates with and without county  $i$  in the estimation.

## 5 Conclusion

We have shown that protesters motivated historically underrepresented groups to run for a seat in Congress. These results suggest that collective actions such as protests can help to improve the representation of women and minorities. Moreover, our findings have at least three implications. First, previous research has shown that changes in the representation of groups in the population leads to policy changes, hence we should expect historically underrepresented groups to benefit from their improved representation. Second, having more Congresswomen elected can potentially help to reduce stereotypes and the negative bias in female leaders' effectiveness. Finally, although we focus on high-profile political positions, the Women's March could have also impacted the private sector and lower rank positions.

## References

- Ahrens, A., Hansen, C., and Schaffer, M. (2018). pdlasso and ivlasso: Programs for post-selection and post-regularization OLS or IV estimation and inference.
- Aidt, T. S. and Franck, R. (2015). Democratization under the threat of revolution: Evidence from the great reform act of 1832. *Econometrica*, 83(2):505–547.
- Bagues, M. and Esteve-Volart, B. (2010). Can gender parity break the glass ceiling? Evidence from a repeated randomized experiment. *Review of Economic Studies*, (77):1301–1328.
- Banducci, S., Donovan, T., and Karp, J. (2004). Minority representation, empowerment, and participation. *Journal of Politics*, 66(2):534–556.
- Baskaran, T. and Hessami, Z. (2018). Does the election of a female leader clear the way for more women in politics? *American Economic Journal: Economic Policy*, (10):95–121.
- Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: Does exposure reduce bias? *Quarterly Journal of Economics*, 124(4):1497–1540.
- Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *Science*, 335:582–586.

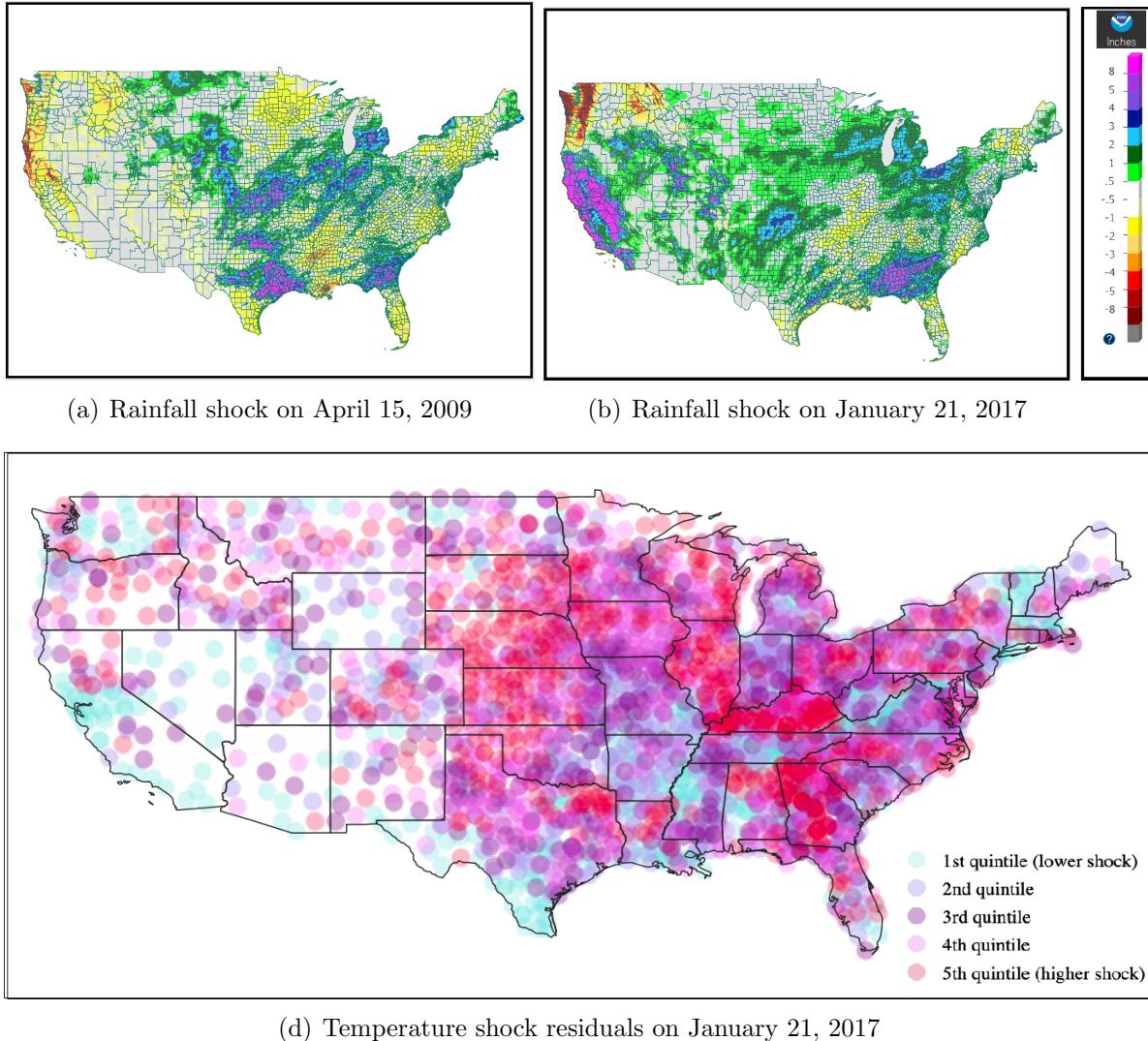
- Belloni, A., Chernozhukov, V., and Hansen., C. (2011). Lasso methods for gaussian instrumental variables models. <https://arxiv.org/abs/1012.1297>.
- Beyerlein, K., Peter, R., Aliyah, A.-H., and Amity., P. (2018). The 2017 women's march: A national study of solidarity events. *Mobilization: An International Quarterly*, 23(4):425–449.
- Bialik, K. (2019). For the fifth time in a row, the new congress is the most racially and ethnically diverse ever. Fact Tank. Pew Research Center.
- Bialik, K. and Krogstad, J. M. (2017). 115th Congress sets new high for racial, ethnic diversity. Fact Tank. Pew Research Center.
- Bond, D., Bond, J., Oh, C., Jenkings, C., and Taylor, C. (2003). Integrated data for events analysis (IDEA): An event typology for automated events data development. *Journal of Peace Research*, 40(6):733–745.
- Bond, D., Jenkings, C., and Taylor, C. (1997). Mapping mass political conflict and civil society: Issues and prospects for the automated development of event data. *Journal of Conflict Resolution*, 41(4):553–579.
- Brollo, F. (2016). What happens when a woman wins an election? Evidence from close races in Brazil. *Journal of Development Economics*, (122):28–45.
- Broockman, D. (2009). Do female politicians empower women to vote or run for office? a regression discontinuity approach. *Electoral Studies*, (34):190–204.
- Button, J. (1978). *Black Violence: Political Impact of the 1960s Riots*. Princeton University Press.
- Chattopadhyay, R. and Duflo, E. (2004). Women as policymakers: Evidence from a randomized policy experiment in india. *Econometrica*, (5):1409–1443.
- Chenoweth, E. and Pressman, J. (2017). This is what we learned by counting the Women's Marches. The Washington Post.
- Chenoweth, E. and Stephan, M. (2012). *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Columbia University Press.
- Collins, W. and Margo, R. (2007). The economic aftermath of the 1960s riots in American cities: Evidence from property values. *Journal of Economic History*, 67(4):849–883.
- Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 91(1):1–45.
- Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly exogenous. *The Review of Economics and Statistics*, 94(1):260–272.
- Dal Bó, E., Finan, F., Folke, O., Persson, T., and Rickne, J. (2017). Who becomes a politician? *Quarterly Journal of Economics*, 132(4):1877–1914.

- Desilver, D. (2018). A record number of women will be serving in the new congress. Fact Tank. Pew Research Center.
- Dittmar, K. (2018). Putting the Record Numbers of Women's Candidacies into Context. Center for American Women and Politics.
- Duflo, E. (2005). Why political reservations? *Journal of the European Economic Association*, 3(2):668–678.
- Dunning, T. and Nilekani, J. (2013). Ethnic quotas and political mobilization: caste, parties, and distribution in Indian village councils. *American Political Science Review*, 107(1):35–56.
- Enos, R., Kaufman, A., and Sands, M. (2019). Can violent protest change local policy support? Evidence from the aftermath of the 1992 Los Angeles riot. *American Political Science Review*.
- Ferreira, F. and Gyourko, J. (2014). Does gender matter for political leadership? The case of u.s. mayors. *Journal of Public Economics*, 112:24–39.
- Fisher, D. R., Andrews, K. T., Caren, N., Chenoweth, E., Heaney, M. T., Leung, T., Perkins, N., and Pressman, J. (2019). The science of contemporary street protest: New efforts in the United States. *Science Advances*, 5(10):eaaw5461.
- Fisher, D. R., Dow, D. M., and Ray., R. (2017). Intersectionality Takes it to the Streets: Mobilizing across Diverse Interests for the Women's March. *Science Advances*, 3(9):1–8.
- Franklin, J. (2009). Contentious challenges and government responses in Latin America. *Political Research Quarterly*, 62(4):700–714.
- Frey, S., Dietz, T., and Kalof, L. (1992). Characteristics of successful American protest groups: Another look at gamson's strategy of social protest. *American Journal of Sociology*, 98(2):368–387.
- Ghani, E., Kerr, W., and O'Connell, S. (2014). Political reservations and women's entrepreneurship in India. *Journal of Development Economics*, (108):138–153.
- Gilardi, F. (2015). The temporary importance of role models for women's political representation. *American Journal of Political Science*, 59:957–970.
- Hill, C. (2018). Women Make Strides in State-Level Primaries in 2018. Center for American Women and Politics.
- Hsiang, S., Burke, M., and Miguel, E. (2013). Quantifying the influence of climate on human violence. *Science*, 341(6151).
- Huet-Vaughn, E. (2013). Quiet riot: The causal effect of protest violence. *Working Paper*.

- Isaac, L. and Kelly, W. (1981). Racial insurgency, the state, and welfare expansion: Local and national level evidence from the postwar United States. *American Journal of Sociology*, 86(6):1348–1386.
- Iyer, L., Mani, A., Mishra, P., and Topalova, P. (2012). The power of political voice: Women's political representation and crime in India. *American Economic Journal: Applied Economics*, 4(4):165–193.
- Madestam, A., Shoag, D., Veuger, S., and Yanagizawa-Drott, D. (2013). Do Political Protests Matter? Evidence from the Tea Party Movement. *Quarterly Journal of Economics*, 128:1633–85.
- McAdam, D. (1982). *Political Process and the Development of Black Insurgency, 1930-1970*. University of Chicago Press.
- McAdam, D. and Su, Y. (2002). The war at home: Antiwar protests and congressional voting, 1965 to 1973. *American Sociological Review*, 67(5):696–721.
- O'Connell, S. (2018). Political inclusion and educational investment: Estimates from a national policy experiment in India. *Journal of Development Economics*, (135):478–487.
- O'Connell, S. (2019). Can quotas increase the supply of candidates for higher-level positions? Evidence from local government in India. *Review of Economics and Statistics*.
- Pande, R. (2003). Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India. *American Economic Review*, 4(93):1132–1151.
- Pettigrew, S. (2018). November 2018 general election results (county-level). Harvard Data-verse. V1.
- Preuhs, R. (2006). Minority representation, empowerment, and participation. *Journal of Politics*, 68(3):585–599.
- Sass, T. and Mehay, S. (2003). Minority representation, election method, and policy influence. *Economics and Politics*, 15(3):323–339.
- Segura, G. and Bowler, S. (2006). *Diversity in Democracy: Minority Representation in the United States*. University of Virginia Press.
- Sheppard, D. and Glassberg, E. (2016). Something to Talk About: Social Spillovers in Movie Consumption. *Journal of Political Economy*, 124(5).
- Shorter, E. and Tilly, C. (1971). Le déclin de la grève violente en france de 1890 à 1935. *Le Mouvement social*, (76):95–118.
- Snyder, D. and Kelly, W. (1976). Industrial violence in Italy, 1878-1903. *American Journal of Sociology*, 82(1):131–162.

- Snyder, J. M. and Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, 118(2):355–408.
- Strömberg, D. (2015). Media and politics. *Annual Review of Economics*, 7:173–205.
- Washington, E. (2012). Do majority-black districts limit blacks' representation? The case of the 1990 redistricting. *Journal of Law and Economics*, 55:251–274.
- Welch, S. (1976). The impact of urban riots on urban expenditures. *American Journal of Political Science*, 19(4):741–760.

Figure 1: Geographic distribution of weather shocks on protest days



Notes: Panels (a) and (b) present the precipitation departures from averages for the months of April 2009 and January 2017, respectively. The bottom panel shows the residuals of the standardized temperature shock on January 21st, 2017. We calculate these residuals after adjusting for a vector of LASSO-chosen and predetermined county characteristics. Images in panels (a) and (b) were obtained from the National Weather Services, Advanced Hydrologic Prediction Service.

Table 1: Descriptive statistics

|   | All              | Counties<br>with<br>protests | Counties<br>without<br>protests | Difference<br>(3)-(2) | Lasso-chosen weather variable |                 |
|---|------------------|------------------------------|---------------------------------|-----------------------|-------------------------------|-----------------|
|   | (1)              | (2)                          | (3)                             | (4)                   | (5)                           | (6)             |
| <i>Demographic characteristics</i>          |                  |                              |                                 |                       |                               |                 |
| Female population (%)                       | 50.77<br>(1.26)  | 50.87                        | 50.67                           | 0.19                  | 0.27<br>(0.05)                | 0.44<br>(0.05)  |
| Foreign-born population (%)                 | 38.30<br>(33.83) | 46.76                        | 29.63                           | 17.13                 | -20.35<br>(3.97)              | -9.63<br>(2.06) |
| African American population (%)             | 12.52<br>(12.75) | 12.24                        | 12.81                           | -0.57                 | 3.95<br>(0.82)                | 6.97<br>(1.02)  |
| Hispanic population (%)                     | 17.72<br>(17.18) | 21.94                        | 13.39                           | 8.56                  | 10.57<br>(1.03)               | 6.50<br>(1.09)  |
| White population (%)                        | 73.14<br>(16.49) | 69.91                        | 76.44                           | -6.54                 | 4.22<br>(1.88)                | 1.71<br>(1.06)  |
| Median household income (log)               | 10.93<br>(0.26)  | 10.97                        | 10.90                           | 0.07                  | 0.06<br>(0.02)                | 0.05<br>(0.02)  |
| Unemployment rate (%)                       | 5.22<br>(1.63)   | 5.07                         | 5.37                            | -0.31                 | 0.05<br>(0.11)                | 0.00<br>(0.00)  |
| Education, less than college (%)            | 69.76<br>(10.77) | 66.79                        | 72.80                           | -6.02                 | 1.52<br>(0.47)                | -0.83<br>(0.35) |
| <i>Electoral characteristics</i>            |                  |                              |                                 |                       |                               |                 |
| Democrat vote share in 2014 (%)             | 45.75<br>(21.10) | 51.07                        | 40.30                           | 10.77                 | -4.99<br>(1.25)               | 0.00<br>(0.00)  |
| Republican vote share in 2014 (%)           | 50.40<br>(20.65) | 44.76                        | 56.18                           | -11.42                | 6.07<br>(1.24)                | 0.00<br>(0.00)  |
| Turnout in 2014 (%)                         | 24.04<br>(7.88)  | 23.37                        | 24.73                           | -1.36                 | 2.04<br>(0.92)                | -0.13<br>(0.17) |
| Hillary Clinton vote share in 2016 (%)      | 48.42<br>(17.03) | 54.91                        | 41.78                           | 13.12                 | -6.10<br>(1.32)               | 1.03<br>(0.22)  |
| Donald Trump vote share in 2016 (%)         | 45.97<br>(17.01) | 39.07                        | 53.05                           | -13.98                | 7.11<br>(1.11)                | 0.00<br>(0.00)  |
| Turnout in 2016 (%)                         | 42.21<br>(7.62)  | 41.75                        | 42.69                           | -0.94                 | 2.21<br>(0.74)                | 0.37<br>(0.14)  |
| Women vote share 2016 (%)                   | 20.25<br>(22.87) | 24.65                        | 15.74                           | 8.90                  | -4.24<br>(1.10)               | -1.70<br>(1.18) |
| Underrepresented groups vote share 2016 (%) | 33.33<br>(28.79) | 38.97                        | 27.55                           | 11.41                 | -8.44<br>(1.50)               | -1.93<br>(1.12) |
| Counties                                    | 2,945            | 471                          | 2,474                           | 2,945                 | 2,945                         | 2,945           |

Notes: Column 1 presents means and standard deviations in parenthesis. Column 2 (3) present means for counties with a positive (zero) number of protesters on January 21st, 2017. All means are weighted by population. Column (4) presents the difference between columns 2 and 3. All differences in column 4 are statistically significant at conventional levels except for African American population, and turnout in both 2014 and 2016. Columns (5) and (6) present the cross-sectional correlation between the lasso-chosen weather variable (i.e. temperature shock) and the corresponding county characteristics with (column 6) and without (column 5) controlling for other county characteristics.

Table 2: The effect of weather shocks on attendance to the Women's March

|                               | Dependent variable: Protesters population (%) |                |                |                 |                 |                 |
|-------------------------------|---|----------------|----------------|-----------------|-----------------|-----------------|
|                               | (1)   | (2)            | (3)            | (4)             | (5)             | (6)             |
| Rainy protest indicator       | 0.21<br>(0.32)                                | 0.16<br>(0.30) |                | -0.42<br>(0.47) |                 |                 |
| Rainfall                      |   |                | 0.09<br>(0.24) |                 |                 |                 |
| LASSO-chosen weather variable |   |                |                |                 | -0.51<br>(0.12) | -0.71<br>(0.12) |
| Counties/Districts            | 2931  | 2931           | 2931           | 464             | 2935            | 329             |
| R-squared                     | 0.275   | 0.246          | 0.275          | 0.413           | 0.144           | 0.391           |
| F-statistic                   | 0.41  | 0.28           | 0.15           | 0.82            | 18.30           | 34.19           |
| Protesters variable           | Best Guess                                    | Low Estimate   | Best Guess     | Best Guess      | Best Guess      | Best Guess      |
| Counties/Districts            | All   | All            | All            | Protesters>0    | All             | All             |
| Election controls             | Y   | Y              | Y              | Y               | N               | N               |
| Demographic controls          | Y   | Y              | Y              | Y               | N               | N               |
| LASSO-chosen controls         | N   | N              | N              | N               | Y               | Y               |
| Avg. dependent variable       | 1.04  | 0.84           | 1.04           | 2.04            | 1.04            | 1.04            |

Note: The unit of analysis is a county in columns 1-5 and congressional districts in column 6. A rainy protest is defined based on the precipitation amount on January 21st, 2017. The rainy protest indicator equals one if there was more than 0.1 inches of rain. Rainfall in column 3 is the precipitation amount in inches. The variable chosen by LASSO is the standardized average temperature shock: January 21st, 2017's average temperature deviation from its mean, divided by its standard deviation. Robust standard errors in parentheses clustered at the state level.

Table 3: The Women’s March, weather shocks, and the 2018 House of Representatives Election

|                                | Candidates      |                 |                                 | Vote shares (%) |                             |                 |     |
|--------------------------------|-----------------|-----------------|---------------------------------|-----------------|-----------------------------|-----------------|-----|
|                                | Total           | Women (%)       | All underrepresented groups (%) | Women           | All underrepresented groups | Turnout (%)     |     |
|                                |                 |                 | (1)                             | (2)             | (3)                         | (4)             | (5) |
| <i>Reduced Form</i>            |                 |                 |                                 |                 |                             |                 |     |
| LASSO-chosen weather variable  | 0.17<br>(0.11)  | -7.66<br>(1.60) | -9.01<br>(2.06)                 | -3.29<br>(1.38) | -5.15<br>(1.20)             | -0.84<br>(0.26) |     |
| <i>Two-stage least squares</i> |                 |                 |                                 |                 |                             |                 |     |
| Protesters (%)                 | -0.25<br>(0.15) | 11.07<br>(3.37) | 16.55<br>(5.75)                 | 9.66<br>(5.48)  | 12.71<br>(5.20)             | 1.55<br>(0.57)  |     |
| <i>Ordinary least squares</i>  |                 |                 |                                 |                 |                             |                 |     |
| Protesters (%)                 | -0.02<br>(0.04) | 2.92<br>(1.22)  | 1.30<br>(1.03)                  | 0.79<br>(0.53)  | 0.25<br>(0.38)              | 0.11<br>(0.06)  |     |
| Counties/Districts             | 328             | 328             | 328                             | 2,935           | 2,935                       | 2,935           |     |
| Avg. dependent variable        | 2.66            | 26.08           | 36.45                           | 27.78           | 41.14                       | 35.02           |     |

Note: All outcomes are measured in the 2018 House of Representatives Election. The LASSO-chosen weather variable is the standardized average temperature shock: January 21st, 2017’s average temperature deviation from its mean, divided by its standard deviation. The outcomes are: total candidates in column 1, the share of candidates that are women in column 2, the ones that belong to an underrepresented group in politics in column 3 – i.e. women, Hispanic, African-American, Asians/Pacific Islanders or Native Americans – the vote shares obtained by women and underrepresented groups in columns 4 and 5, and turnout in the same election in column 6. The unit of analysis is a congressional district in columns 1-3 and a county in columns 4-6. All regressions are population weighted and include LASSO-chosen controls for each specification. Robust standard errors in parentheses, clustered at the state level.

Table 4: Results by underrepresented group

|                                | Women candidates (%) |                 |                  |                      | Women vote share (%) |                 |                  |                      |
|--------------------------------|----------------------|-----------------|------------------|----------------------|----------------------|-----------------|------------------|----------------------|
|                                | Hispanic             | Non Hispanic    | African American | Non African American | Hispanic             | Non Hispanic    | African American | Non African American |
|                                | (1)                  | (2)             | (3)              | (4)                  | (5)                  | (6)             | (7)              | (8)                  |
| <i>Reduced form</i>            |                      |                 |                  |                      |                      |                 |                  |                      |
| LASSO-chosen weather variable  | -0.65<br>(0.49)      | -6.95<br>(1.63) | 1.62<br>(0.59)   | -9.22<br>(1.41)      | -1.12<br>(0.48)      | -4.68<br>(1.27) | 2.13<br>(0.83)   | -6.06<br>(1.38)      |
| <i>Two-stage least squares</i> |                      |                 |                  |                      |                      |                 |                  |                      |
| Protesters (%)                 | -                    | 10.05<br>(3.21) | -2.33<br>(1.01)  | 13.32<br>(3.60)      | 1.36<br>(0.85)       | 7.89<br>(2.28)  | -5.02<br>(2.24)  | 14.26<br>(4.85)      |
| <i>Ordinary least squares</i>  |                      |                 |                  |                      |                      |                 |                  |                      |
| Protesters (%)                 | 0.92<br>(0.86)       | 2.35<br>(1.46)  | -0.33<br>(0.33)  | 3.60<br>(1.04)       | -0.05<br>(0.13)      | 1.12<br>(0.49)  | 0.13<br>(0.38)   | 0.64<br>(0.46)       |
| Counties/Districts             | 328                  | 328             | 328              | 328                  | 2,935                | 2,935           | 2,935            | 2,935                |
| Avg. dependent variable        | 2.09                 | 23.96           | 2.94             | 23.11                | 2.25                 | 25.54           | 4.93             | 22.85                |

Note: All outcomes are measured in the 2018 House of Representatives Election. The LASSO-chosen weather variable is the standardized average temperature shock: January 21st, 2017's average temperature deviation from its mean, divided by its standard deviation. The outcomes are: the share of candidates that are Hispanic women in columns 1 and 5, non-Hispanic women in columns 2 and 6, African American women in columns 3 and 7, and non-African American women in columns 4 and 8. The unit of analysis is a congressional district in columns 1-4 and a county in columns 5-8. All regressions are population weighted and include LASSO-chosen controls for each specification. Robust standard errors in parentheses, clustered at the state level.

# Online Appendix

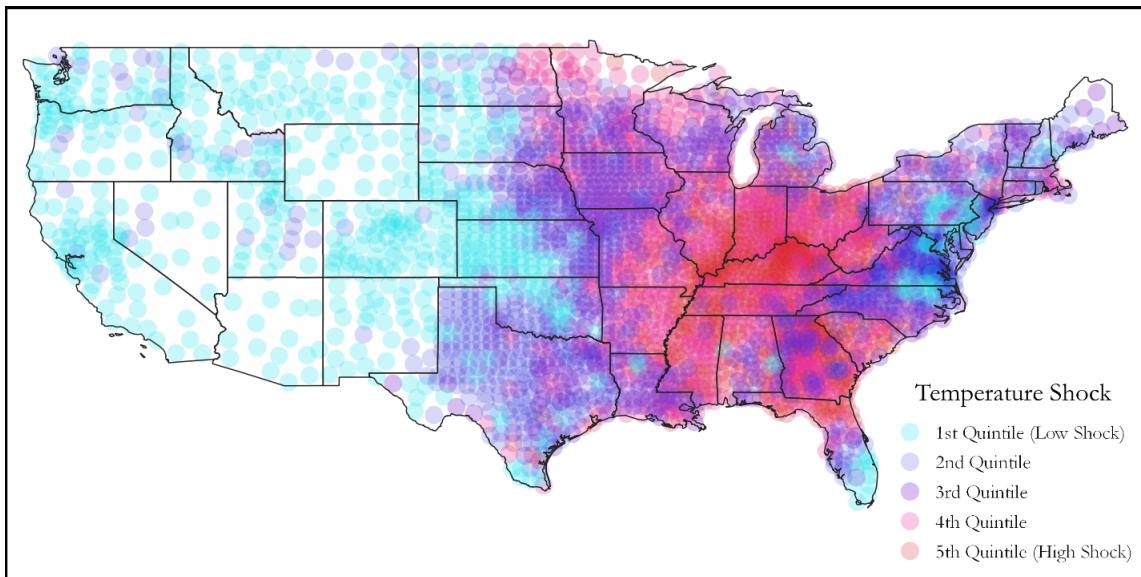
## List of Figures

|     |   |     |
|-----|---|-----|
| A.1 | Temperature shock without residualizing . . . . . | ii  |
| A.2 | Non-parametric first stage . . . . .              | iii |
| A.3 | Plausible exogeneity test . . . . .               | iv  |
| A.4 | Robustness of two-stage estimates . . . . .       | v   |
| A.5 | Robustness of first-stage . . . . .               | vi  |

## List of Tables

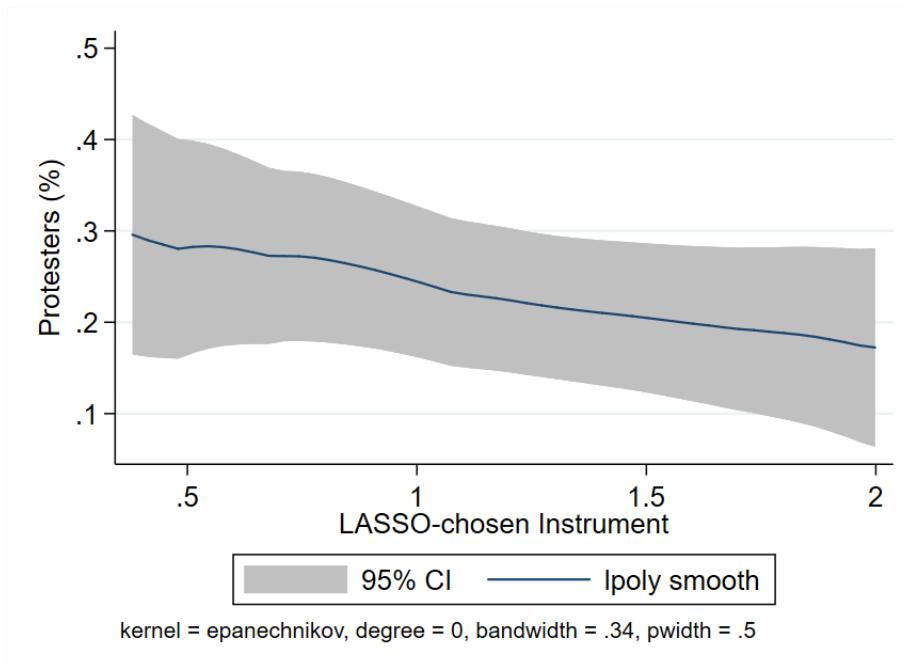
|      |   |      |
|------|---|------|
| A.1  | Descriptive statistics for districts in and out of the sample . . . . .               | vii  |
| A.3  | Vector of weather shocks - possible instruments . . . . .                             | ix   |
| A.4  | Vector of possible controls . . . . .   | x    |
| A.5  | Machine-chosen controls . . . . .   | xi   |
| A.6  | Alternative specifications for the first-stage . . . . .                              | xii  |
| A.7  | Robustness of results to spatial correlation . . . . .                                | xiii |
| A.8  | Robustness of results to human-selected controls . . . . .                            | xiv  |
| A.9  | Local reports of protesters in counties with <i>high</i> temperature shocks . . . . . | xv   |
| A.10 | Local reports in counties with <i>low</i> temperature shocks . . . . .                | xvi  |
| A.11 | Robustness of 2SLS Results to excluding outliers based on their DFBETA .              | xvii |

Figure A.1: Temperature shock without residualizing



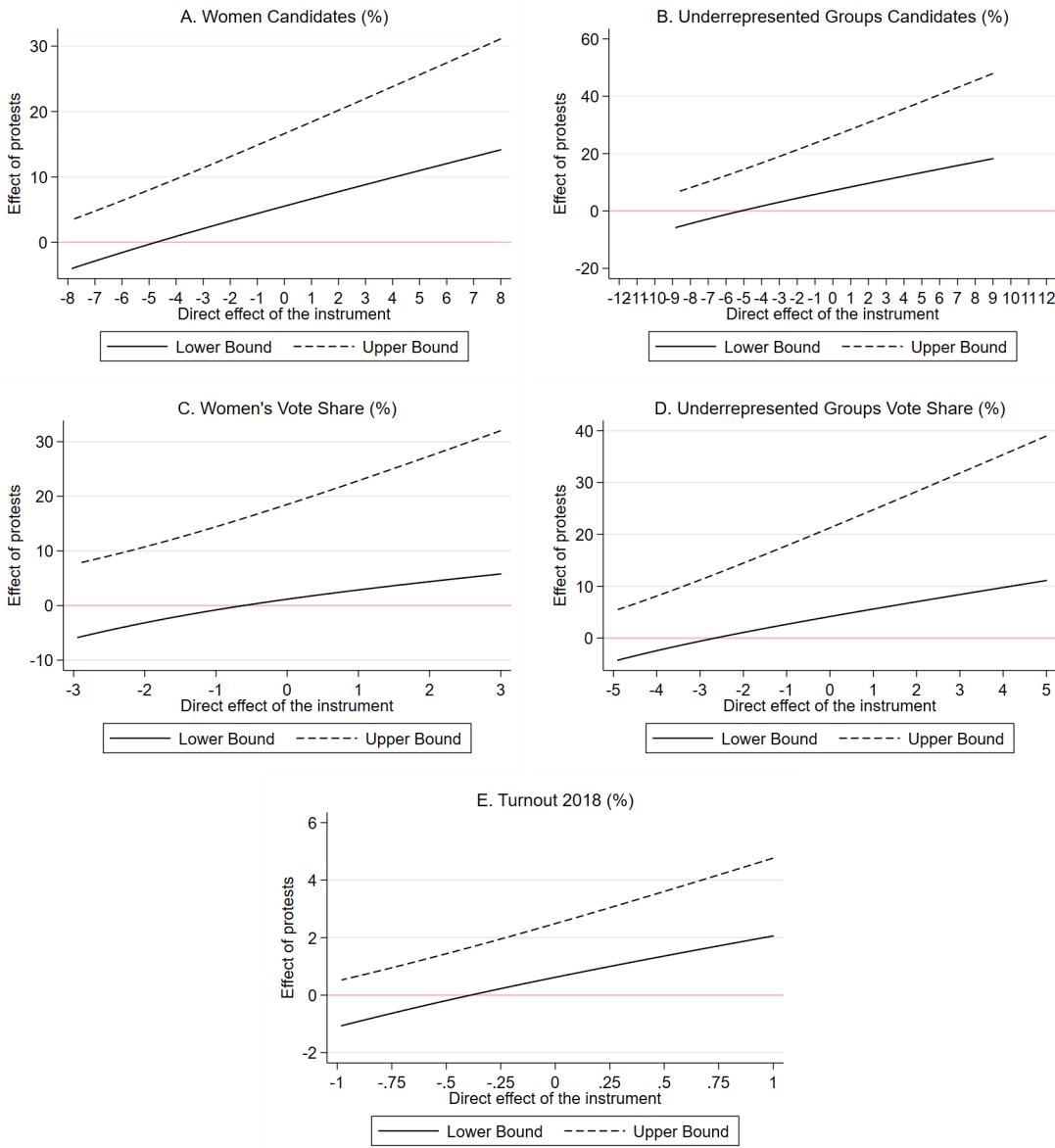
Notes: Geographic distribution of temperature shocks on January 21, 2017. This shock is defined as  $z_i \equiv \frac{x_i - \bar{x}_i}{\sigma_i}$ , where  $x_i$  is the average temperature in county  $i$  the day of the Women's March and  $\bar{x}_i, \sigma_i$  are the average and standard deviation of  $x_i$  calculated using five random days in January during the seven years before the March.

Figure A.2: Non-parametric first stage



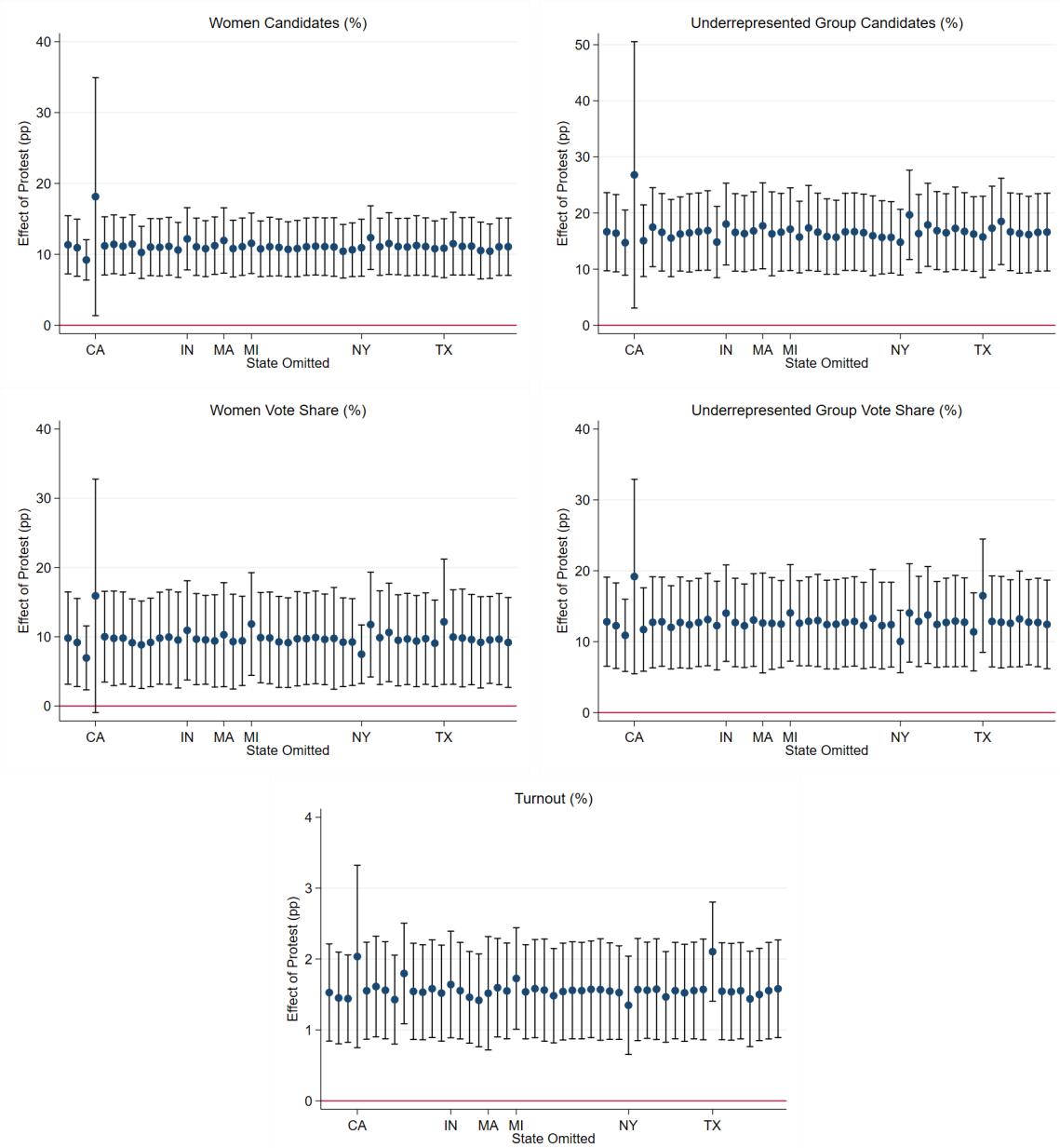
Notes: Protesters (%) is the share of protesters per capita in a county. The instrument chosen by LASSO is the Standardized Average Temperature Shock: January 21st, 2017's average temperature deviation from its mean, divided by its standard deviation. The sample is in the county-level and includes the observations between the 10th and 90th percentiles of the instrument.

Figure A.3: Plausible exogeneity test



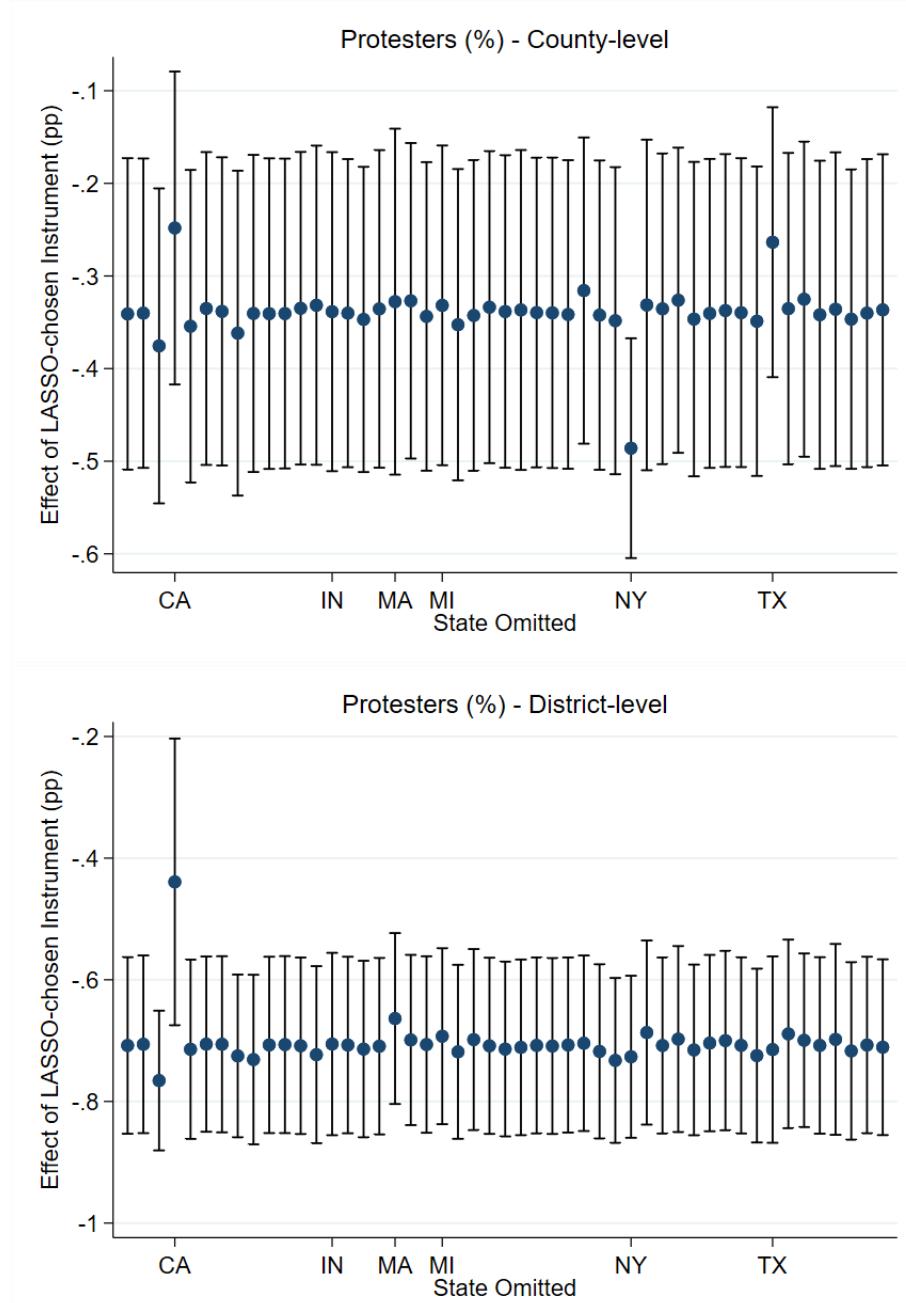
Notes: These figures present results from a bounding exercise in which we allow the temperature shock to affect outcomes directly. The x-axis measures (theoretical) direct effects of temperature shock on Women Candidates (Panel A), Underrepresented Group Candidates (Panel B), Women's vote share (Panel C), Underrepresented Groups' vote share (Panel D) and Turnout (Panel E). The y-axis measures the corresponding effect of protests. Overall, we find that to make the effect of protests non-different from zero the direct effect of the instrument would have to be -4.6 in panel A, -5.1 in Panel B, -0.6 in Panel C, -2.7 in Panel D and -0.4 in Panel E, equivalent to 60% (-4.6/-7.6), 57% (-5.1/-9.0), 18% (-0.6/-3.29), 52% (-2.7/-5.15) and 48% (-0.4/-0.84) of the reduced form effects.

Figure A.4: Robustness of two-stage estimates



Notes: Figure A.4 presents the variation in the results of Table III, Panel B, when omitting one state at a time. Underrepresented Group includes Women, Hispanics, African-Americans, Asians/Pacific Islanders and Native Americans.

Figure A.5: Robustness of first-stage



Notes: Figure A.5 presents the variation in the First Stage results, when omitting one state at a time. Underrepresented Group includes Women, Hispanics, African-Americans, Asians/Pacific Islanders and Native Americans.

Table A.1: Descriptive statistics for districts in and out of the sample

|                                    | All<br>(1)       | In<br>(2) | Out<br>(3) | Difference<br>(4) |
|------------------------------------|------------------|-----------|------------|-------------------|
| <i>Demographic characteristics</i> |                  |           |            |                   |
| Female population (%)              | 50.77<br>(1.26)  | 50.46     | 50.99      | -0.53             |
| Foreign-born population (%)        | 38.30<br>(33.83) | 17.32     | 52.83      | -35.50            |
| African American population (%)    | 12.52<br>(12.75) | 9.73      | 14.45      | -4.72             |
| Hispanic population (%)            | 17.71<br>(17.18) | 10.67     | 22.60      | -11.93            |
| White population (%)               | 73.13<br>(16.49) | 81.86     | 67.09      | 14.78             |
| Median household income (log)      | 10.93<br>(0.26)  | 10.83     | 11.01      | -0.18             |
| Unemployment rate (%)              | 5.22<br>(1.63)   | 5.50      | 5.02       | 0.47              |
| Education, less than college (%)   | 69.76<br>(10.77) | 74.93     | 66.18      | 8.75              |
| <i>Electoral characteristics</i>   |                  |           |            |                   |
| Democrat vote share 2014 (%)       | 45.75<br>(21.10) | 35.42     | 52.82      | -17.40            |
| Republican vote share 2014 (%)     | 50.40<br>(20.65) | 60.61     | 43.41      | 17.20             |
| Turnout 2014                       | 24.04<br>(7.88)  | 26.73     | 22.18      | 4.55              |
| Hillary Clinton vote share (%)     | 48.42<br>(17.03) | 37.13     | 56.24      | -19.11            |
| Donald Trump vote share (%)        | 45.97<br>(17.01) | 57.25     | 38.17      | 19.08             |
| Turnout in 2016 (%)                | 42.21<br>(7.62)  | 43.87     | 41.06      | 2.81              |
| Counties                           | 2,946            | 2,558     | 388        | 2,946             |

Notes: Column (1) presents means and standard deviations for all counties. Column (2) presents means for counties that are subsets of Congressional Districts and Column (3) for counties that are not. Column (4) presents the mean difference coefficients between both and their respective significance. Means are population weighted. Standard errors are clustered at the state level. \*\*\* 1% , \*\* 5% , \* 10% significance

Table A.2: Descriptive statistics for additional demographic characteristics

|                                   | All<br>(1)       | Protest<br>(2) | No Protest<br>(3) | Difference<br>(4) |
|-----------------------------------|------------------|----------------|-------------------|-------------------|
| Family households (%)             | 66.42<br>(6.09)  | 64.89          | 67.98             | -3.10***          |
| Female citizens (%)               | 36.43<br>(3.33)  | 35.98          | 36.88             | -0.90*            |
| Unmarried partners households (%) | 6.11<br>(1.29)   | 6.44           | 5.77              | 0.68***           |
| Rural population (%)              | 18.77<br>(24.05) | 9.36           | 28.41             | -19.05***         |
| Unemployment change (2013-2017)   | -1.47<br>(1.28)  | -1.70          | -1.23             | -0.47***          |
| Population density (logs)         | 6.13<br>(1.62)   | 6.52           | 5.72              | 0.80***           |
| Aged under 30 (%)                 | 39.84<br>(4.47)  | 40.36          | 39.30             | 1.06***           |
| Aged over 65 (%)                  | 14.50<br>(3.81)  | 14.02          | 14.99             | -0.97***          |
| Number of counties                | 2,945            | 471            | 2,474             | 2,945             |

Notes: Column 1 presents means and standard deviations in parenthesis. Column 2 (3) present means for counties with a positive (zero) number of protesters on January 21st, 2017. All means are weighted by population. Column (4) presents the difference between columns 2 and 3. All differences in column 4 are statistically significant at conventional levels.

Table A.3: Vector of weather shocks - possible instruments

| Description                                      | Average Temperature               | Maximum Temperature               | Rain                              |
|--|-----------------------------------|-----------------------------------|-----------------------------------|
| Deviation from historical mean                   | Shock                             | Shock                             | Shock                             |
| Squared shock                                    | Squared shock                     | Squared shock                     | Squared shock                     |
| Cubed shock                                      | Cubed shock                       | Cubed shock                       | Cubed shock                       |
| Shock divided by historical standard deviation   | Standardized shock                | Standardized shock                | Standardized shock                |
| Squared shock divided by historical sd           | Squared shock standardized        | Squared shock standardized        | Squared shock standardized        |
| Absolute value of shock divided by historical sd | Absolute value shock standardized | Absolute value shock standardized | Absolute value shock standardized |
| Shock bins                                       | Shock bins (1-5)                  | Shock bins (1-6)                  | Shock bins (1-16)                 |
| Dummy for each bin                               | 5 2F shock bins                   | 6 2F shock bins                   | 16 0.25 inches rain shock bins    |
| Indicator for any rain                           |                                   |                                   | Any rain                          |
| Indicator for any snow                           |                                   |                                   | Any snow                          |

Table A.4: Vector of possible controls

| Demographic                       | Electoral                           |
|-----------------------------------|-------------------------------------|
| Female population (%)             | Clinton vote share                  |
| Family households (%)             | Trump vote share                    |
| Foreign-born population (%)       | Votes for Clinton (% o population)  |
| Median household income (log)     | Votes for Trump (% o population)    |
| Unemployment rate (%)             | Turnout 2016                        |
| Unemployment change (2013-2017)   | Democratic Party vote share (2014)  |
| African American population (%)   | Republican Party vote share (2014)  |
| Hispanic population (%)           | Votes for DP 2014 (% of population) |
| Population density (log)          | Votes for RP 2014 (% of population) |
| Rural population (%)              | Turnout 2014                        |
| White population (%)              |                                     |
| Female citizens (%)               |                                     |
| Unmarried partners households (%) |                                     |
| Distance to Washington DC (log)   |                                     |
| 10 deciles of population dummies  |                                     |

Table A.5: Machine-chosen controls

|                                | LASSO-chosen controls  | Number of controls not chosen |
|--------------------------------|--|-------------------------------|
| <i>District-level analysis</i> |  |                               |
| Total Candidates               | Democratic Party Vote Share (2014), Share of Family Households (%), Unemployment Rate (%)  | 31                            |
| Women                          | Democratic Party Vote Share (2014), Share of Family Households (%), Unemployment Rate (%)  | 31                            |
| Underrepresented groups        | Democratic Party Vote Share (2014), Share of Family Households (%), Unemployment Rate (%), Trump Vote Share  | 30                            |
| <i>County-level analysis</i>   |  |                               |
| Women                          | Trump Vote Share, Votes for Trump (% of population), Democratic Party Vote Share (2014), Republican Party Vote Share (2014), Votes for DP 2014 (% of population), Unemployment Rate (%), Second Decile Population, Ninth Decile Population | 26                            |
| Underrepresented groups        | Clinton Vote Share, Votes for Trump (% of population), Democratic Party Vote Share (2014), Republican Party Vote Share (2014), Votes for DP 2014 (% of population), Unemployment Rate (%), Ninth Decile Population                         | 27                            |
| Turnout 2018 (%)               | Turnout 2016, Turnout 2014, Democratic Party Vote Share (2014), Republican Party Vote Share (2014), Votes for DP 2014 (% of population), Unemployment Rate (%), First Decile Population, Ninth Decile Population                           | 26                            |

Notes: The flexible controls for population size are dummies for each decile on the variable's distribution (i.e. Second Decile Population is an indicator for having a low share of population, corresponding to the second decile in the population size distribution.)

Table A.6: Alternative specifications for the first-stage

|                                | Protesters (%)  |                 |                 | Protesters (thousands) |                   |                   | Log protesters  |
|--------------------------------|-----------------|-----------------|-----------------|------------------------|-------------------|-------------------|-----------------|
|                                | (1)             | (2)             | (3)             | (4)                    | (5)               | (6)               | (7)             |
| <b>County-level</b>            |                 |                 |                 |                        |                   |                   |                 |
| LASSO-chosen weather variable  | -0.51<br>(0.12) | -0.23<br>(0.08) | -0.11<br>(0.45) | -0.23<br>(0.08)        | -41.73<br>(13.15) | -39.79<br>(11.03) | -0.46<br>(0.12) |
| Counties/Districts             | 2935            | 2935            | 468             | 2935                   | 2935              | 468               | 440             |
| F-statistic                    | 18.3            | 7.4             | 0.1             | 7.6                    | 10.1              | 13.0              | 15.8            |
| Protesters                     | Best Guess      | Low Estimate    | Best Guess      | Best Guess             | Low Estimate      | Best Guess        | Best Guess      |
| Sample                         | All             | All             | Protesters>0    | All                    | All               | Protesters>0      | Protesters>0    |
| LASSO-chosen controls          | Y               | Y               | Y               | Y                      | Y                 | Y                 | Y               |
| Avg. dependent variable        | 1.04            | 0.84            | 2.04            | 1.07                   | 0.86              | 6.77              | 6.14            |
| <b>Congressional Districts</b> |                 |                 |                 |                        |                   |                   |                 |
| LASSO-chosen weather variable  | -0.71<br>(0.12) | -0.38<br>(0.08) | -0.63<br>(0.12) | -55.50<br>(13.05)      | -27.70<br>(3.43)  | -55.56<br>(12.87) | -1.34<br>(0.09) |
| Counties/Districts             | 329             | 329             | 232             | 330                    | 330               | 233               | 233             |
| F-statistic                    | 34.2            | 24.5            | 30.1            | 18.1                   | 65.1              | 18.6              | 207.5           |
| Protesters                     | Best Guess      | Low Estimate    | Best Guess      | Best Guess             | Low Estimate      | Best Guess        | Best Guess      |
| Sample                         | All             | All             | Protesters>0    | All                    | All               | Protesters>0      | Protesters>0    |
| LASSO-chosen Controls          | Y               | Y               | Y               | Y                      | Y                 | Y                 | Y               |
| Avg. dependent variable        | 1.04            | 0.84            | 1.33            | 10.18                  | 8.22              | 14.41             | 7.58            |

Note: The unit of analysis is a county in Panel A and a Congressional District in Panel B. The instrument chosen by LASSO is the Standardized Average Temperature Shock: January 21st, 2017's average temperature deviation from its mean, divided by its standard deviation. Controls are also LASSO-chosen, and are mainly composed by previous electoral outcomes, flexible dummies for population and measures of unemployment. Best Guess denotes the average turnout across the three estimations of attendance data. Low estimate is the derived most conservative count of the turnout in any given location. Regressions in columns 1-3 are population weighted. Robust standard errors in parentheses, clustered at the state level. \*\*\* 1% , \*\* 5% , \* 10% significance.

Table A.7: Robustness of results to spatial correlation

|                                 | Candidates (%)  |                             | Vote shares (%) |                             |                |
|---------------------------------|-----------------|-----------------------------|-----------------|-----------------------------|----------------|
|                                 | Women           | All underrepresented groups | Women           | All underrepresented groups | Turnout (%)    |
|                                 |                 | (1)                         |                 | (2)                         |                |
| <i>Distance cutoff: 100 kms</i> |                 |                             |                 |                             |                |
| Protesters (%)                  | 11.07<br>(3.77) | 16.55<br>(9.05)             | 9.66<br>(5.25)  | 12.71<br>(6.28)             | 1.55<br>(0.51) |
| <i>Distance cutoff: 50 kms</i>  |                 |                             |                 |                             |                |
| Protesters (%)                  | 11.07<br>(3.71) | 16.55<br>(9.80)             | 9.66<br>(7.07)  | 12.71<br>(8.09)             | 1.55<br>(0.57) |
| Counties/Districts              | 328             | 328                         | 2935            | 2935                        | 2935           |
| Avg. dependent variable         | 26.1            | 36.5                        | 27.8            | 41.1                        | 35.0           |

Note: This table shows the effect of Protests, instrumented with a LASSO-chosen instrument, on the Electoral Outcomes with standard errors adjusted for spatial correlation, as proposed by Conley (1999), using Collela et al. (2019)'s program. I use distance cutoffs for the spatial kernel of 100kms in Panel A and 50kms in Panel B. All Underrepresented Groups includes Women, Hispanic, African-American, Asians/Pacific Islanders and Native Americans. The unit of analysis is a congressional district in columns 1 and 2 and a county in columns 3-5. Temperature Shock is the Standardized Average Temperature Shock: January 21st, 2017's average temperature deviation from its mean, divided by its standard deviation. Both the mean and standard deviation are constructed with historical observations for each weather station. All regressions are population weighted and include LASSO-chosen controls for each specification, which are mainly composed by previous electoral outcomes, flexible dummies for population and measures of unemployment. Robust standard errors in parentheses, adjusted for spatial correlation. \*\*\* 1% , \*\* 5% , \* 10% significance

Table A.8: Robustness of results to human-selected controls

|                                | Candidates (%)  |                             | Vote shares (%) |                             |                 |
|--------------------------------|-----------------|-----------------------------|-----------------|-----------------------------|-----------------|
|                                | Women           | All underrepresented groups | Women           | All underrepresented groups | Turnout (%)     |
|                                |                 | (1)                         |                 | (2)                         |                 |
| <i>Reduced Form</i>            |                 |                             |                 |                             |                 |
| LASSO-Chosen weather variable  | -6.35<br>(3.23) | -12.20<br>(3.14)            | -3.67<br>(1.61) | -5.61<br>(1.66)             | -0.73<br>(0.29) |
| <i>Two-stage least squares</i> |                 |                             |                 |                             |                 |
| Protesters (%)                 | 8.22<br>(4.44)  | 15.80<br>(6.10)             | 9.92<br>(6.46)  | 15.15<br>(8.22)             | 1.98<br>(1.36)  |
| <i>Ordinary least squares</i>  |                 |                             |                 |                             |                 |
| Protesters (%)                 | 2.61<br>(0.78)  | 1.66<br>(0.70)              | 0.57<br>(0.46)  | 0.43<br>(0.38)              | 0.01<br>(0.06)  |
| Counties/Districts             | 328             | 328                         | 2,935           | 2,935                       | 2,935           |
| Avg. dependent variable        | 26.08           | 36.45                       | 27.78           | 41.14                       | 35.02           |

Note: LASSO-Chosen weather variable is a temperature shock on January 21, 2017. The outcomes are the share of women candidates and the share of candidates from underrepresented group in politics: Women, Hispanic, African-American, Asians/Pacific Islanders or Native Americans. Vote shares and turnout for the 2018 House of Representatives Election. The unit of analysis is a congressional district in columns 1 and 2 and a county in columns 3-5. All regressions are population weighted and include the same controls as in Madestam et al. (2013) plus a vector of women-related controls. Robust standard errors in parentheses, clustered at the state level.

Table A.9: Local reports of protesters in counties with *high* temperature shocks

| County ID | Value of the instrument | Protesters (%) | Local report | Local newspaper                |
|-----------|-------------------------|----------------|--------------|--------------------------------|
| (1)       | (2)                     | (3)            | (4)          | (5)                            |
| 12001     | 2,51                    | 0,76           | Y            | The Gainsville Sun             |
| 12073     | 2,05                    | 5,50           | Y            | Tallahassee Democrat           |
| 17019     | 2,17                    | 2,56           | Y            | The News Gazette               |
| 17031     | 2,06                    | 4,78           | Y            | Chicago Tribune                |
| 17077     | 2,05                    | 3,22           | Y            | The Southern Illinoisan        |
| 17089     | 2,06                    | 0,11           | N            | –                              |
| 17143     | 2,09                    | 0,93           | Y            | WMBD News                      |
| 18003     | 2,27                    | 0,27           | Y            | The Journal Gazette            |
| 18097     | 2,31                    | 0,71           | Y            | Indiana Public Media           |
| 18127     | 2,11                    | 0,22           | Y            | The Times of Northwest Indiana |
| 18157     | 2,22                    | 0,47           | Y            | Journal and Courier            |
| 18167     | 2,29                    | 0,18           | Y            | Tribune Star                   |
| 21035     | 2,09                    | 1,81           | Y            | WKMS                           |
| 21067     | 2,14                    | 2,27           | Y            | WKYT                           |
| 21111     | 2,04                    | 0,65           | Y            | Courier Journal                |
| 26077     | 2,32                    | 0,56           | Y            | M Live                         |
| 26161     | 2,06                    | 3,34           | Y            | Ground Cover News              |
| 39035     | 2,19                    | 1,20           | Y            | Cleveland.com                  |
| 39095     | 2,10                    | 0,05           | Y            | The Blade                      |
| 42049     | 2,31                    | 1,15           | Y            | Goerie.com                     |
| 45077     | 2,10                    | 0,41           | Y            | Independent Mail               |
| 47157     | 2,03                    | 0,61           | Y            | Memphis Flyer                  |

Notes: Own construction.

Table A.10: Local reports in counties with *low* temperature shocks

| County ID<br>(1) | Value of the instrument<br>(2) | Protesters (%)<br>(3) | Local report<br>(4) | Local newspaper<br>(5)    |
|------------------|--------------------------------|-----------------------|---------------------|---------------------------|
| 4005             | -0,64                          | 1,75                  | Y                   | Arion Daily Sun           |
| 4019             | -1,33                          | 1,55                  | Y                   | Tucson.com                |
| 6007             | -1,34                          | 0,84                  | Y                   | Chico Enterprise Record   |
| 6013             | -0,90                          | 0,54                  | Y                   | San Francisco Chronicle   |
| 6027             | -0,51                          | 3,33                  | Y                   | Bronco Roundup            |
| 6037             | -1,07                          | 4,45                  | Y                   | Los Angeles Times         |
| 6055             | -0,78                          | 2,12                  | Y                   | Napa Valley Register      |
| 6057             | -1,43                          | 0,25                  | Y                   | The Union                 |
| 6061             | -1,64                          | 0,17                  | Y                   | Tahoe Daily Tribune       |
| 6073             | -0,99                          | 1,23                  | Y                   | KPBS                      |
| 6079             | -0,92                          | 2,96                  | Y                   | The Tribune               |
| 6083             | -0,66                          | 1,56                  | Y                   | Santa Barbara Independent |
| 6085             | -1,20                          | 1,64                  | Y                   | San Francisco Chronicle   |
| 6087             | -0,37                          | 4,19                  | Y                   | Santa Cruz Sentinel       |
| 6111             | -0,90                          | 0,27                  | N                   | –                         |
| 15009            | -1,27                          | 1,88                  | Y                   | The Maui News             |
| 30049            | -0,33                          | 14,97                 | Y                   | Independent Record        |
| 49053            | -0,81                          | 0,83                  | Y                   | St. George News           |
| 53005            | -0,41                          | 0,87                  | Y                   | Tri-City Herald           |
| 53031            | -0,78                          | 1,99                  | Y                   | Peninsula Daily News      |
| 53071            | -0,38                          | 3,63                  | Y                   | KEPR                      |

Notes: Own construction.

Table A.11: Robustness of 2SLS Results to excluding outliers based on their DFBETA

|                         | Candidates (%)  |                             | Vote shares (%) |                             |                |
|-------------------------|-----------------|-----------------------------|-----------------|-----------------------------|----------------|
|                         | Women           | All Underrepresented groups | Women           | All Underrepresented groups | Turnout (%)    |
|                         |                 | (1)                         |                 | (2)                         |                |
| Protesters (%)          | 12.61<br>(3.29) | 16.61<br>(4.88)             | 9.41<br>(3.93)  | 12.34<br>(4.72)             | 1.34<br>(0.51) |
| Districts/Counties      | 308             | 308                         | 2,755           | 2,755                       | 2,755          |
| Avg. dependent variable | 26.08           | 36.45                       | 27.78           | 41.14                       | 35.02          |

Note: This table shows the effect of protesters, instrumented with a LASSO-chosen instrument, on the Electoral Outcomes when excluding observations based on their DFBETA. Following the standard approach, I exclude all observations for which  $|DFBETA_i| < \frac{2}{\sqrt{(N)}}$  where N is the number of observations. All Underrepresented Groups includes Women, Hispanic, African-

American, Asians/Pacific Islanders and Native Americans. The unit of analysis is a congressional district in columns 1 and 2 and a county in columns 3-5. Temperature Shock is the Standardized Average Temperature Shock: January 21st, 2017's average temperature deviation from its mean, divided by its standard deviation. Both the mean and standard deviation are constructed with historical observations for each weather station. All regressions are population weighted and include LASSO-chosen controls for each specification, which are mainly composed by previous electoral outcomes, flexible dummies for population and measures of unemployment. Robust standard errors in parentheses, clustered at the state level. \*\*\* 1% , \*\* 5% , \* 10% significance