INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

560
2021

# The Impact of the Women's March on the U.S. House Election

**Felipe González y Magdalena Larreboure.**

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

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

Magdalena Larreboure          Felipe González

March 8, 2021

Three million people participated in the Women's March against discrimination in 2017, the largest single-day protest in U.S. history. We show that 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 increased turnout at the Election and the vote shares obtained by minorities. We conclude that protests can help to empower historically underrepresented groups.

*Keywords*: protests, election, gender, minority

*We would like to thank Álvaro Cordero, Emilio Depetris-Chauvin, Katia Everke, Sebastián Figari, 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. Larreboure: Busara Center for Behavioral Economics, Nairobi, Kenya. González: 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 in the public sphere. In this paper we estimate the local impact of protesters on women and people from ethnic minorities who ran 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.[1]

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 turnout at the Election and the vote share of candidates from historically underrepresented groups.

We begin by replicating Madestam et al.'s (2013) econometric strategy, who use an indicator for 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 differences in the geographic distribution of rainfall between the day of the Tea Party protest and the day of the Women's March, or different motives behind these protests, among others. Building on their strategy and the work of Gilchrist and Sands (2016), we create a vector with dozens of weather shocks and choose the best predictors of local protest attendance using the least absolute shrinkage and selection 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 empirically a strong predictor of attendance to the March. Importantly, this temperature shock is conditionally uncorrelated with previous political outcomes and it is also uncorrelated with protest participation when measured in previous years.

[1]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.

Using the machine-chosen weather shock as an instrument for the local intensity of the Women's March in a two-stage least squares framework, we find that protesters increased the vote share of women and other candidates from ethnic minorities. More precisely, we estimate that 1,000 additional protesters – the observed size of the average protest in a county – increased the vote share of women and minorities by approximately 13 percentage points (3,000) more votes in a county, close to 32% of the sample mean. Remarkably, most of this change in voting patterns is explained by an increase in the vote for white women (2,500 votes) with the remaining increase being explained by additional votes for man from ethnic minorities. Notably, women candidates from ethnic minorities do *not* appear to systematically benefit from the political impacts of the March.

Our analysis addresses empirical issues related to omitted variables and measurement error in local protests. However, we still need to face the possibility of a potential violation of the exclusion restriction. A leading concern mentioned in previous literature 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. Most importantly, we followed Gentzkow and Shapiro (2010) and constructed a new county-level dataset with media articles covering the March and found that the number of articles is unrelated to the temperature shock we exploit. Moreover, our findings are also robust to omitting 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 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 add to a growing literature studying underrepresented groups and ways to improve their representation. Our main contribution is to show that collective actions such as protests can empower these groups by pushing 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, 2020). 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 an empirical literature that estimates the economic and political impacts of protests and other collective actions. 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, women's position within households, and property values (Collins and Margo, 2007; Madestam et al., 2013; Aidt and Franck, 2015; Bargain et al., 2019; González and Vial, 2021).[2] 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 in the public sphere.

## 2 The Context

The Women's March took place on January 21st 2017 and it was a massive event.[3] Although the beginning of this movement was tied to the election of the Republican Donald Trump as President, surveys reveal that 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. In fact, underrepresented groups different from women constitute 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, 2017). Representation has improved but it is still far from matching the U.S. population.

In terms of representing the U.S. population, 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

[2] 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, 2020; 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).

[3] 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).

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 was the number of congresswomen elected the highest in U.S. history, it was also the biggest jump in women members since the 1990s. This can be easily seen in the fact that 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. This multi-sourced 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 (Fisher et al., 2019). Because the CCC reports are originally at the city level, we aggregated these to the county level to match the outcomes we examine. Most cities belong to a single county, hence this aggregation was straightforward, and when this was not the case we assigned the city to the county with the largest share. Reports were pulled together if more than one city protested within a county.

We downloaded the weather data from the National Oceanic and Atmospheric Administration (NOAA). In particular, 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. The large vector of weather variables we use pushes us to drop some stations with incomplete weather data. Besides a wide vector 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.[4] 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,

[4]We use average and maximum temperature and exclude minimum temperatures because they usually occur during the night and protests take place during the day.

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 main 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 three outcome variables. (i) the vote shares obtained by women, (ii) the vote share obtained by candidates from underrepresented groups, and (iii) turnout. The underrepresented groups in this study include women, African-Americans, Hispanics, Asian/Pacific Islanders, and Native Americans. To determine whether candidates represented minority groups, we use data from The Asian Pacific American Institute for Congressional Studies (APAICS), blackwomeninpolitics.com, NALEO Educational Fund ("Election 2018 Races to Watch: The Power of Latino Candidates"), and "History, Art & Archives, U.S. House of Representatives." When needed, we complement this information with data from the candidates' websites.

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 white population, a larger share of foreign born and Hispanic population, and host more educated people with higher median income and less unemployment. Politically, counties with and without protests have similar turnout, but the former are more Democrat and voted relatively more for women and other underrepresented groups in the previous election. 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 $i$, $Protesters_i$ 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:

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

where $Protesters_i$ is a measure of attendance to the march in county $i$. $Rain_i$ 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. $Likelihood\ of\ Rain_i$ 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 capita. Standard errors are clustered at the state level, but results are robust 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 $\widehat{\beta}$ 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 Gilchrist and Sands (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:

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

where $WeatherShock_i$ are the LASSO-chosen instruments. The chosen variable is the standardized temperature shock the day of the March.[5] Figure 1 presents a map with the variation of this shock after removing the variation from the vector of machine-chosen control variables.[6]

[5]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 $i$ 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.1 presents the vector with all possible weather shocks to be chosen.

[6]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

Importantly, the machine-chosen weather shock has little empirical relationship with previous electoral variables. Columns 5 and 6 in Table 1 present estimates of equation (3) using county characteristics as dependent variable. To avoid cherry-picking $w_i$, these are also LASSO-chosen, but the results are similar if we use a fixed set of control variables. The estimates in these columns reveal that $w_i$ is important because (i) all electoral differences across counties disappear after including $w_i$ in the estimation (panel B), and (ii) the weather shock affected counties with less foreign population, more African Americans, and more Hispanics (panel A). Therefore all specifications will include machine-chosen controls for each dependent variable.[7]

# 4 Results

## *4.1 Attendance to the Women's March*

Table 2 presents estimates of equations (2) and (3) to test for the impact of weather shocks on attendance to the March. 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. Columns 1, 3, and 4 measure the number of protesters using an estimate from a variety of sources – what the CCC reports call "Best guess" – and column 2 uses the lowest reported number ("Low estimate").

The results indicate that rainfall the day of the event has little predictive power on the size of local protests. If anything, the sign of the relationship is the opposite of what we expected. We highlight three 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. Consistent with this explanation, panels A and B 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 (Beyerlein et al., 2018). 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. And third, different weather shocks might be important for turnout decisions in January versus April.

In contrast to the rainfall shock, the machine-chosen temperature shock has a strong predictive power on protest participation (see Table 2, column 5). The results in this column 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., $0.51 \times 0.84 = 0.43$). This coefficient

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.

[7] There are 24 socio-economic and 10 electoral predetermined variables to be potentially chosen as controls. Table A.2 presents all of these and Table A.3 shows the set chosen for each outcome. Results are similar if we use the controls employed by Madestam et al. (2013).

represents a 43% change with respect to the sample average and the associated $F$-statistic is 17.[8] Moreover, in the appendix we show that the temperature shocks in January 21 of previous years (2011-2016) are empirically unrelated with the number of protesters in 2017, with inconsistent coefficients that are smaller in magnitude and sometimes positive or negative (Table A.5).

Why is protest attendance lower with unusually large temperatures? Our interpretation is that the relative price of participating in a protest increases with warmer temperatures during the winter. High temperatures presumably make protesting less attractive because of an increase in the opportunity cost of alternative outdoor activities. Although there is little direct evidence of substitution within the set of outdoor activities, there is some indirect evidence consistent with this notion. In particular, outdoor recreational activities such as biking, running, calisthenics, golf, gardening, and walking increase with warmer temperatures (Graff Zivin and Neidell, 2014), presumably crowding out protest activities. Moreover, this increase in recreational activities is particularly important during winter times (Obradovich and Fowler, 2017; Chan and Wichman, 2020).

### *4.2 The impact of the Women's March*

Table 3 presents the main results of our analysis. Panel A shows the direct effect of the machine-chosen instrument on the outcomes of interest, candidates' vote shares (columns 1-2), and county turnout (column 3).[9] Columns 4-7 present estimates to further understand the impact on subgroup of candidates. Panel B uses the instrument in a two-stage least squares framework to estimate the impact of protesters, and panel C shows OLS results for comparison purposes.[10] Panel A indicates that a one standard deviation increase in the temperature shock on January 21st (0.84) decreased women's vote share and the vote share of underrepresented groups decreased by 4 pp., and turnout decreased by 0.7 pp. In terms of magnitude, each of these estimates represent changes of 18%, 13% and 2% of the sample means respectively. Columns 4-7 show that most of these effects come from changes in the vote shares of Non-Hispanic and Non-African American women.

Two-stage least squares estimates in panel B indicate that the Women's March had an impact on the electoral outcomes of underrepresented groups. To gauge their magnitude, let us consider an increase of 1 pp. in the share of protesters in a county, i.e. approximately 1,000 more protesters which represents the size of the average protest. According to these estimates, protesters increased the vote share of underrepresented groups by 13 pp. (3,000 votes) in the average county. Most

[8]Table A.4 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.

[9]Table A.5 complements the reduced-form results by showing the lack of a relationship between the temperature shock in January 21 of previous years and the electoral outcomes we examine.

[10]Table A.7 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.

of this increase is explained by an increase in the vote share of women, who got 10 percentage points (2,500) more votes. Column 3 shows that protests also motivated citizens to vote, increasing turnout by 1.5 pp. or 1,500 votes. Panel C reveals that a naïve OLS estimation delivers an attenuated coefficient which could be explained by classical measurement error in the number of protesters, omitted variables, or the characteristics of the compliers. Again, in counties with more protesters the additional political support for underrepresented groups favored predominantly white women, with a smaller increase for African-American men and Hispanic men.

Given that the Women's March and the House Election were separated by more than 22 months, what could explain the impact of protesters? Similar to the case of the Tea Party Movement (Skocpol and Williamson, 2016), sustained organizing activity at the local level seems to have played a crucial role. In fact, Putnam and Skocpol (2018) found evidence in local interviews that lead them to conclude that college-educated white women were key: "[W]hat is underway is a national pattern of mutually energizing local engagement. Sociologically, what we are witnessing is an inflection point – a shift in long-standing trends – concentrated in one large demographic group, as college-educated women have ramped up their political participation *en masse*."

### *4.3 The role of the media*

So far our analysis assumes that unusual weather on January 21 of 2017 affected the 2018 election only through attendance to the Women's March. Unusual weather can, however, also change media coverage, which in turn affects electoral outcomes (Snyder and Strömberg, 2010; Strömberg, 2015). We argue that the media is unlikely to threaten our results for several reasons. In the first place, media coverage of the March should be less affected by unusual weather than past protests because of its contextual relevance and the rise of the internet.[11] In this sense, the fact that rainfall has little impact on attendance is reassuring of the March's importance. In the second place, we manually investigated media coverage of the March in counties with weather shocks above the 90th percentile and below the 10th percentile and found media reports for all protests but two (Tables A.8 and A.9). And most importantly, we collected new data on media articles about the Women's March and found these to be uncorrelated with the temperature shock.

To construct the county-level dataset with the number of articles covering the March, we first download all articles available in ProQuest which mentioned the Women's March during the three months before and after January 21 of 2017. Then, we follow Gentzkow and Shapiro (2010) to match each article to a county, dropping foreign articles. Last, we count the number of articles in each county before and after the March. Overall, we found almost 8,000 articles coming from 224 counties, 20% of which were written before January 21 and 80% after that date.

[11] According to surveys conducted by the Pew Research Center, the percentage of Americans with access to the internet has increased from less than 50% in 2000 to 90% in 2020 (Pew Research Center, 2019).

We test for the relationship between the temperature shock and the number of articles using a cross-sectional dataset at the county-level. Columns 8 and 9 in panel A of Table 3 show the lack of a relationship between these variables. Empirically, we estimate equation (3) using as dependent variable the number of articles in the county, the controls from the main specification, and add a control for the number of articles before January 21 of 2017. In particular, column 8 uses as dependent variable an indicator that takes the value of one if there was at least one article covering the Women's March in the three months after and zero otherwise, while column 9 follows Burbidge et al. (1988) and uses the hyperbolic sine transformation of the number of articles. The estimates reveal the lack of a relationship between both variables, with both point estimates being virtually zero and confidence intervals that reject relatively small changes in media coverage.

### *4.4 Other explanations and robustness*

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). Additionally, violence could affect the protesting experience or its effectiveness at the eyes of the general public. 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 18, 47, and 49% of the reduced form effects for the main outcomes. 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.[12]

Lastly, the results are robust to the exclusion of groups of counties.[13] First, the March's impact is 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 impacts of the March are also robust to the exclusion of outliers. To implement this exercise we omit from the estimation 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.

[12]Figure A.3 provides more details about this exercise and the full set of results.

[13]Figure A.4 presents the robustness of two-stage least squares estimates and, for completion, Figure A.5 presents the first-stage. Table A.10 shows estimates omitting outliers.

# 5 Conclusion

We have shown that protesters can empower historically underrepresented groups and improve their political representation significantly. These results suggest that collective actions such as protests have impacts beyond party lines and in that sense complement previous literature. 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). pdslasso 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.

Bargain, O., Boutin, D., and Champeaux, H. (2019). Women's political participation and intra-household empowerment: Evidence from the Egyptian Arab Spring. *Journal of Development Economics*, 141:102379.

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.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401):123–127.

Button, J. (1978). *Black Violence: Political Impact of the 1960s Riots*. Princeton University Press.

Chan, N. W. and Wichman, C. J. (2020). Climate change and recreation: Evidence from North American cycling. *Environmental and Resource Economics*, 76:119–151.

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*, 113(4):1012–1028.

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.

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? evidence from U.S. daily newspapers. *Econometrica*, 78(1):35–71.

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.

Gilchrist, D. S. and Sands, E. G. (2016). Something to talk about: Social spillovers in movie consumption. *Journal of Political Economy*, 124(5):1339–1382.

González, F. and Vial, F. (2021). Collective action and policy implementation: Evidence from Salvador Allende's expropriations. *Journal of Economic History*, 81(2).

Graff Zivin, J. and Neidell, M. (2014). Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics*, 32(1):1–26.

Hsiang, S., Burke, M., and Miguel, E. (2013). Quantifying the influence of climate on human violence. *Science*, 341(6151).

Huet-Vaughn, E. (2020). 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.

Obradovich, N. and Fowler, J. H. (2017). Climate change may alter human physical activity patterns. *Nature Human Behaviour*, 1:97.

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. (2020). Can quotas increase the supply of candidates for higher-level positions? Evidence from local government in India. *Review of Economics and Statistics*, 102(1):65–78.

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 Dataverse. V1.

Pew Research Center (2019). Internet/broadband fact sheet.

Preuhs, R. (2006). Minority representation, empowerment, and participation. *Journal of Politics*, 68(3):585–599.

Putnam, L. and Skocpol, T. (2018). Middle America reboots democracy. *Democracy Journal*.

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.

Shorter, E. and Tilly, C. (1971). Le déclin de la grève violente en france de 1890 à 1935. *Le Mouvement social*, (76):95–118.

Skocpol, T. and Williamson, V. (2016). *The Tea Party and the Remaking of Republican Conservatism*. Oxford University Press.

Snyder, D. and Kelly, W. (1976). Industrial violence in Italy, 1878-1903. *American Journal of Sociology*, 82(1):131–162.

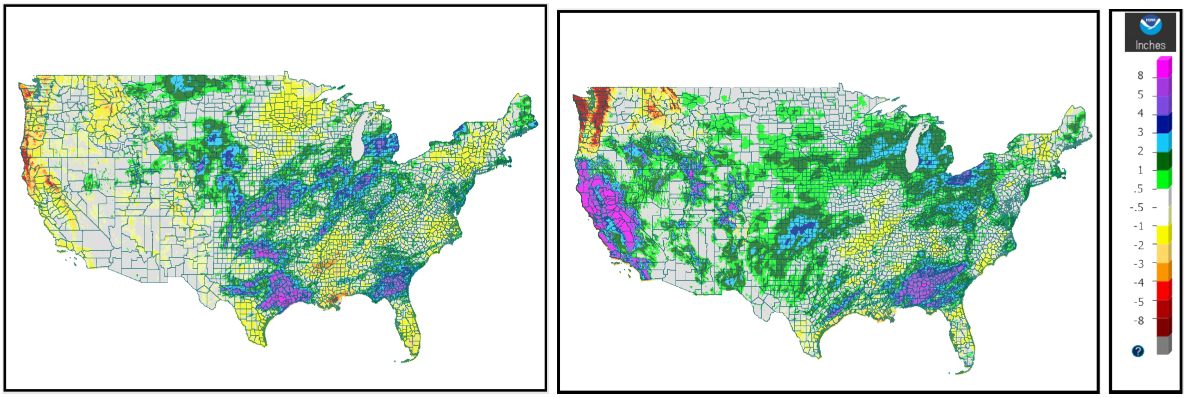
Snyder, J. M. and Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, 118(2):355–408.

Strömberg, D. (2015). Media and politics. *Annual Review of Economics*, 7:173–205.

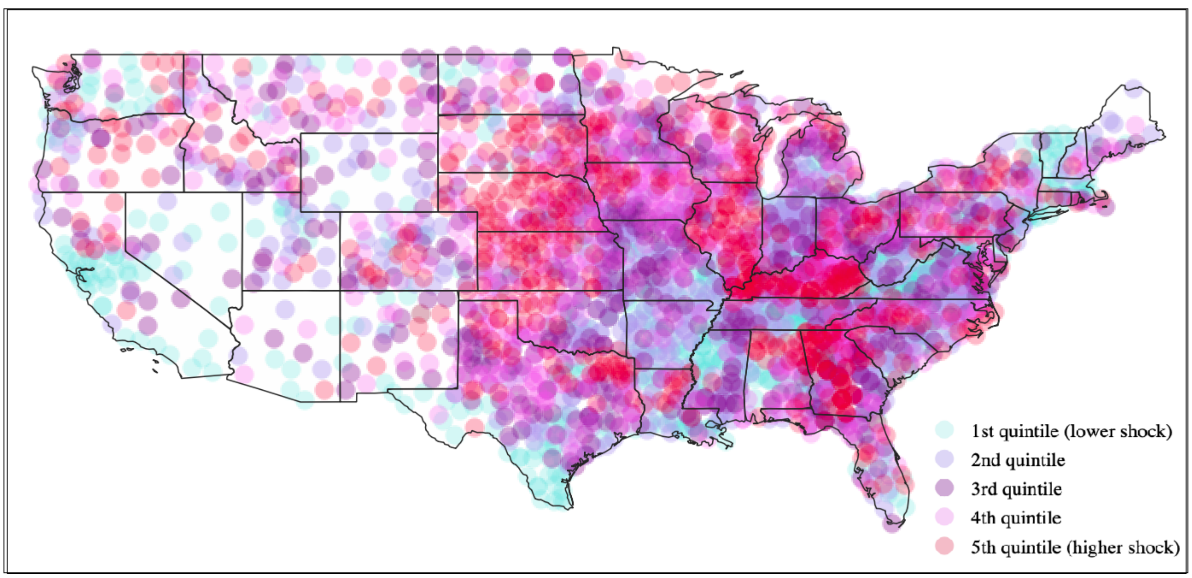
Washington, E. (2012). Do majority-black districts limit blacks' representation? The case of the 1990 redistricting. *Journal of Law and Economics*, 55:251–274.

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




(a) Rainfall shock on April 15, 2009 (b) Rainfall shock on January 21, 2017




(d) Temperature shock residuals on January 21, 2017

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 with protests | Counties without protests | Difference (3)-(2) | Lasso-chosen weather variable: Unconditional exogeneity | Lasso-chosen weather variable: Conditional exogeneity |
|---|---|---|---|---|---|---|
| *Panel A – Demographic characteristics* | (1) | (2) | (3) | (4) | (5) | (6) |
| Female population (%) | 50.77 | 50.87 | 50.67 | 0.19 | 0.27 | 0.41 |
| | (1.26) | | | | (0.05) | (0.06) |
| Foreign-born population (%) | 38.37 | 46.87 | 29.67 | 17.19 | −20.36 | 0.00 |
| | (33.89) | | | | (3.95) | (0.00) |
| African American population (%) | 12.54 | 12.25 | 12.84 | −0.59 | 3.96 | 5.39 |
| | (12.77) | | | | (0.82) | (0.81) |
| Hispanic population (%) | 17.74 | 21.99 | 13.39 | 8.60 | −10.58 | −4.38 |
| | (17.22) | | | | (1.02) | (1.25) |
| White population (%) | 73.12 | 69.90 | 76.41 | −6.51 | 4.21 | −1.36 |
| | (16.50) | | | | (1.88) | (0.86) |
| Median household income (log) | 10.93 | 10.97 | 10.90 | 0.07 | −0.06 | 0.00 |
| | (0.26) | | | | (0.02) | (0.02) |
| Unemployment rate (%) | 5.26 | 5.12 | 5.41 | −0.29 | 0.03 | −0.00 |
| | (1.66) | | | | (0.10) | (0.00) |
| Education, less than college (%) | 69.75 | 66.79 | 72.78 | −6.00 | 1.50 | −2.02 |
| | (10.78) | | | | (0.47) | (0.44) |
| *Panel B – Electoral characteristics* | | | | | | |
| Democrat vote share in 2014 (%) | 45.74 | 51.06 | 40.30 | 10.77 | −4.99 | 0.00 |
| | (21.10) | | | | (1.25) | (0.00) |
| Republican vote share in 2014 (%) | 50.41 | 44.77 | 56.18 | −11.41 | 6.07 | 0.00 |
| | (20.66) | | | | (1.24) | (0.00) |
| Turnout in 2014 (%) | 24.13 | 23.45 | 24.83 | −1.38 | 2.06 | 0.09 |
| | (7.74) | | | | (0.93) | (0.29) |
| Hillary Clinton vote share in 2016 (%) | 48.48 | 54.98 | 41.82 | 13.17 | −6.10 | 1.39 |
| | (17.04) | | | | (1.31) | (0.30) |
| Donald Trump vote share in 2016 (%) | 45.92 | 38.99 | 53.01 | −14.01 | 7.10 | −0.00 |
| | (17.02) | | | | (1.10) | (0.00) |
| Turnout in 2016 (%) | 42.21 | 41.75 | 42.67 | −0.92 | 2.20 | 0.68 |
| | (7.63) | | | | (0.74) | (0.55) |
| Women vote share 2016 (%) | 20.29 | 24.66 | 15.80 | 8.86 | −4.22 | −0.74 |
| | (22.88) | | | | (1.09) | (1.28) |
| Underrepresented groups vote share 2016 (%) | 33.40 | 38.99 | 27.65 | 11.34 | −8.41 | −1.74 |
| | (28.79) | | | | (1.49) | (1.44) |
| Counties | 2,940 | 470 | 2,470 | 2,940 | 2,940 | 2,940 |

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 of which 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 standard errors presented in parentheses; column 6 presents the unconditional correlation and column 5 conditional on 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) |
| Rainy protest indicator | 0.19 | 0.14 | | −0.42 | |
| | (0.29) | (0.27) | | (0.46) | |
| Rainfall | | | −0.04 | | |
| | | | (0.19) | | |
| LASSO-chosen weather variable | | | | | −0.51 |
| | | | | | (0.12) |
| Counties | 2,936 | 2,936 | 2,936 | 466 | 2,940 |
| R-squared | 0.246 | 0.216 | 0.246 | 0.384 | 0.132 |
| F-Statistic | 0.40 | 0.28 | 0.05 | 0.84 | 17.07 |
| Protesters Variable | Best Guess | Low Estimate | Best Guess | Best Guess | Best Guess |
| Counties | All | All | All | Protesters>0 | All |
| Election controls | Y | Y | Y | Y | N |
| Demographic controls | Y | Y | Y | Y | N |
| LASSO-chosen controls | N | N | N | N | Y |
| Avg. dependent variable | 1.00 | 0.79 | 1.00 | 1.98 | 1.00 |

Note: The unit of analysis is a county. 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 and the 2018 House Election

| | Vote shares (%) | | | Women vote share (%) | | | | Media articles about the Women's March | |
|---|---|---|---|---|---|---|---|---|---|
| | Women | All underrepresented groups | Turnout (%) | Hispanic | Non Hispanic | African American | Non African American | Indicator | Logs |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| *Panel A – Reduced Form* | | | | | | | | | |
| LASSO-Chosen weather variable | -4.96 | -5.32 | -0.81 | -0.74 | -4.12 | 2.09 | -6.04 | 0.00 | 0.01 |
| | (1.28) | (1.29) | (0.27) | (0.63) | (1.13) | (0.81) | (1.42) | (0.02) | (0.06) |
| *Panel B – Two-stage least squares* | | | | | | | | | |
| Protesters (%) | 9.73 | 12.95 | 1.52 | 1.32 | 7.47 | -5.00 | 14.42 | – | – |
| | (3.49) | (5.63) | (0.56) | (1.12) | (2.47) | (2.30) | (4.98) | | |
| *Panel C – Ordinary least squares* | | | | | | | | | |
| Protesters (%) | 0.98 | 0.19 | 0.10 | -0.03 | 0.89 | 0.13 | 0.56 | – | – |
| | (0.51) | (0.35) | (0.06) | (0.13) | (0.44) | (0.35) | (0.42) | | |
| Counties | 2,940 | 2,940 | 2,940 | 2,940 | 2,940 | 2,940 | 2,940 | 2,940 | 2,940 |
| Avg. dependent variable | 27.90 | 41.30 | 35.02 | 2.26 | 25.64 | 4.95 | 22.95 | 0.04 | 0.07 |
| Machine-chosen controls | X | X | X | X | X | X | X | X | X |

Note: The outcomes in columns 1-7 are measured in the 2018 House Election. The LASSO-chosen weather variable is the standardized average temperature shock: January 21 of 2017's average temperature deviation from its mean, divided by its standard deviation. The outcomes are: the vote shares obtained by women in column 1, and by candidates that belong to an underrepresented group in politics in column 2 – i.e. women, Hispanic, African-American, Asians/Pacific Islanders or Native Americans – and turnout in the same election in column 3. Columns 4-7 split women's vote shares in mutually exclusive categories of Hispanic and Non-Hispanic women (columns 4-5) and African American and Non-African American women. The outcomes in columns 8 and 9 are an indicator for the appearance of at least one media article (column 8) and the (hyperbolic sine transformation of the) total number of media articles (column 9) about the Women's March in the following three months. The unit of analysis is always a county. All regressions are population weighted and include LASSO-chosen controls for each specification. Standard errors in parentheses are clustered at the state level.

# Online Appendix

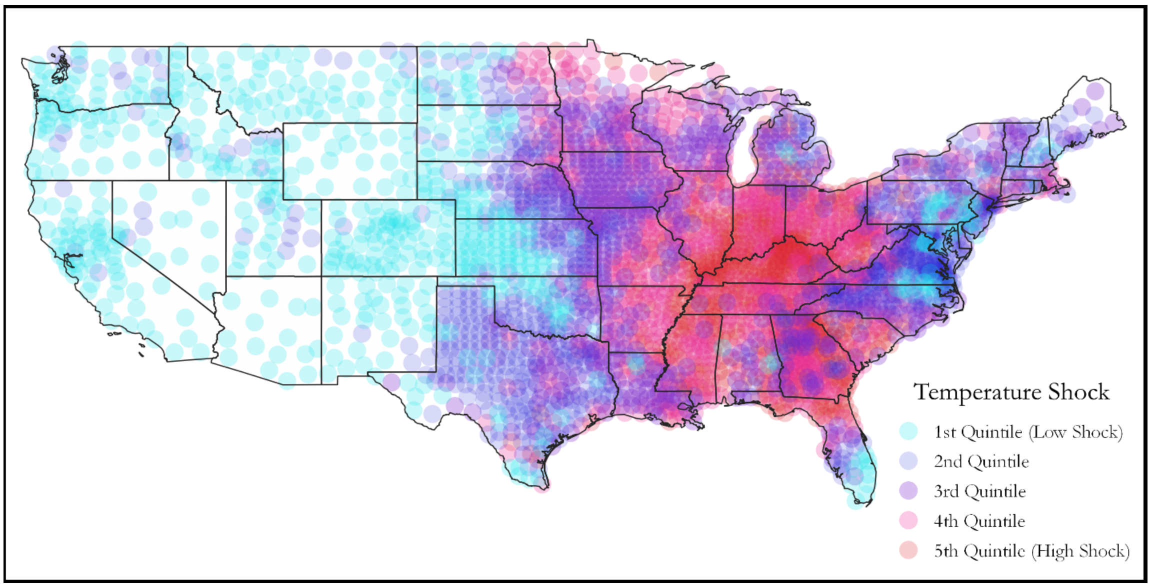
*The Impact of the Women's March on the U.S. House Election*

## List of Figures



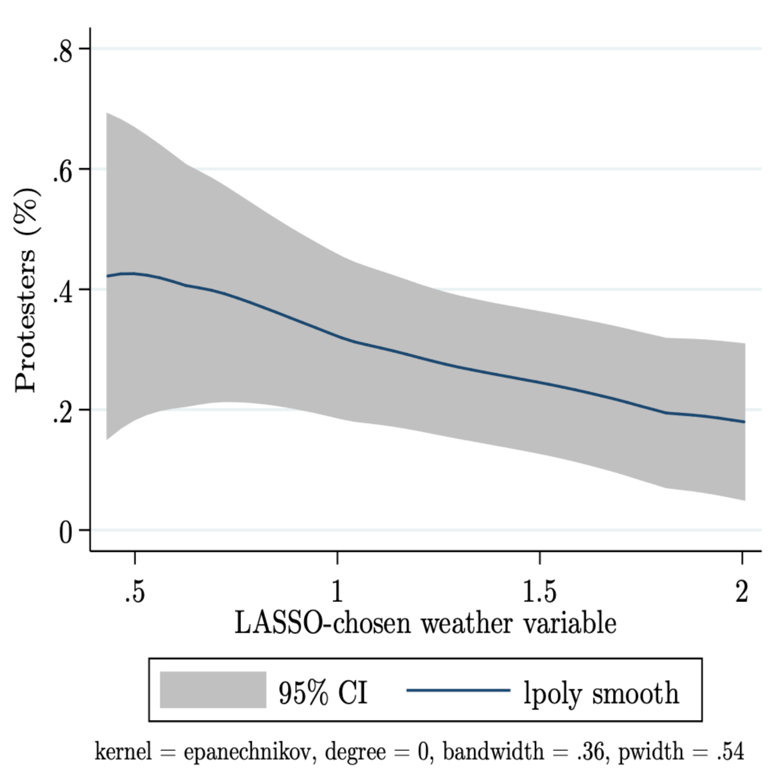
## List of Tables

Figure A.1: Temperature shock without residualizing



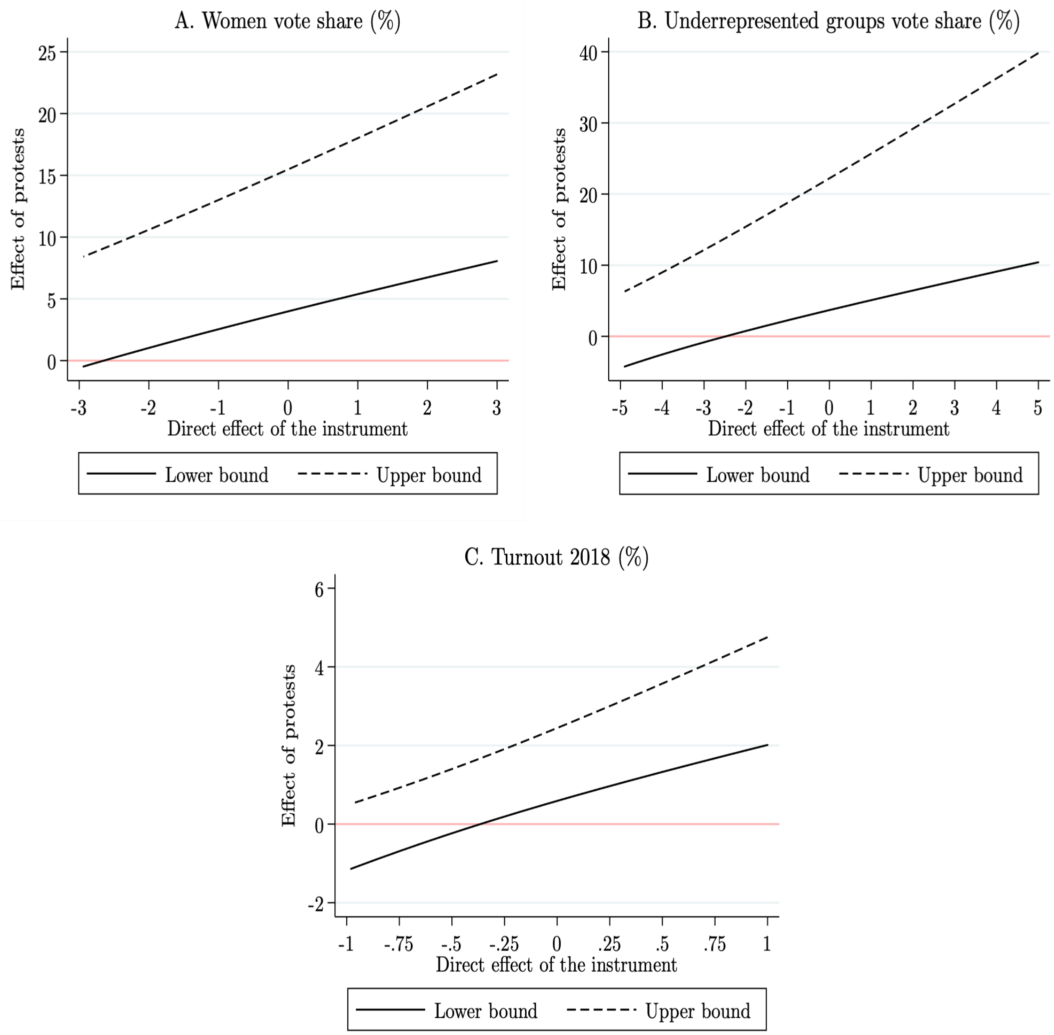

Notes: Geographic distribution of temperature shocks on January 21, 2017. This shock is defined as $z_i \equiv \frac{x_i - \bar{x}_i}{\sigma_i}$, where $x_i$ is the average temperature in county $i$ the day of the Women's March and $\bar{x}_i, \sigma_i$ are the average and standard deviation of $x_i$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

Figure A.2: Non-parametric first stage



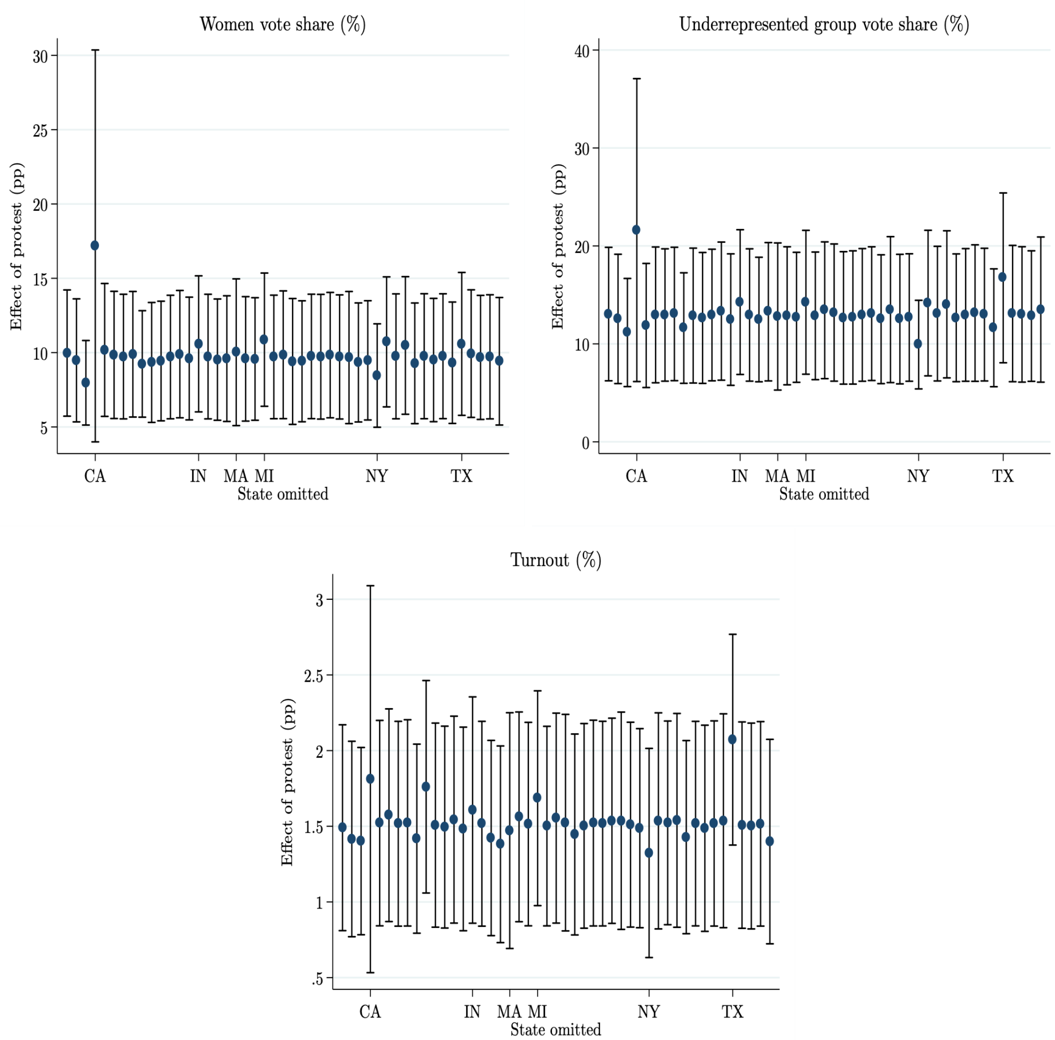

Notes: Non-parametric estimates from a local polynomial estimation. The $y$-axis “Protesters (%)” is the share of protesters per capita in a county. The $x$-axis is the Lasso-chosen instrument, i.e. the average temperature in January 21 of 2017 minus the mean in previous years, divided by its standard deviation. The counties included in this analysis are all of which experienced a temperature shock (the instrument) between the 10th and 90th percentiles of the instrument.

Figure A.3: Plausible exogeneity test



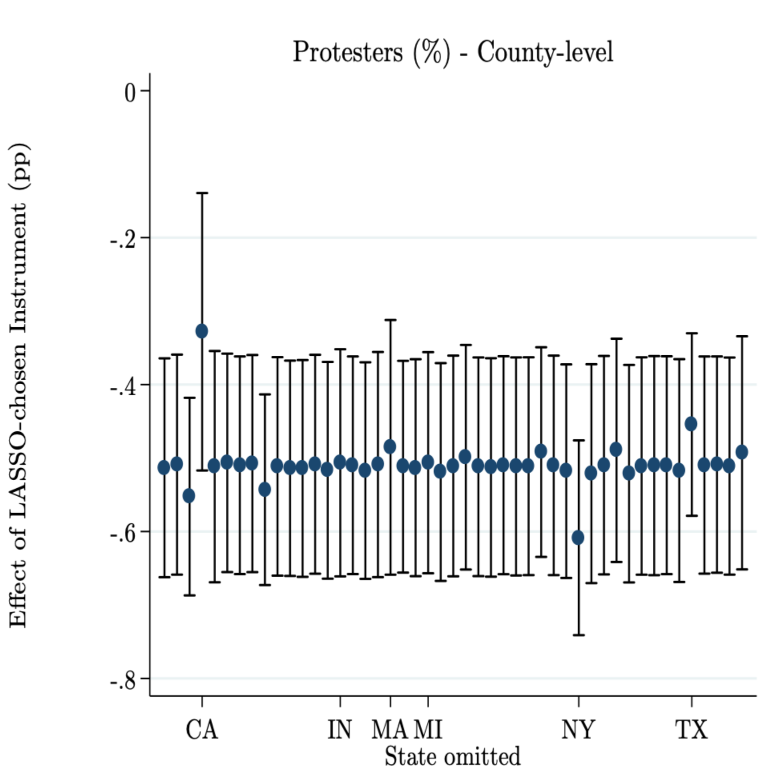

Notes: These figures present results from a bounding exercise in which we allow the temperature shock to affect outcomes directly. The x-axis measures (theoretical) direct effects of temperature shock on women's vote share (Panel A), underrepresented groups' vote share (Panel B) and Turnout (Panel C). 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 -2.6 in Panel A, -2.5 in Panel D and -0.4 in Panel E, equivalent to 18% (-2.6/-4.96), 47% (-2.5/-5.32) and 49% (-0.4/-0.81) of the reduced form effects.

Figure A.4: Robustness of two-stage estimates




Notes: Figure A.4 presents 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 First Stage results, when omitting one state at a time.

Table A.1: 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.2: Vector of possible controls

| Demographic | Electoral |
|---|---|
| Female population (%) | Clinton vote share |
| Family households (%) | Trump vote share |
| Foreign-born population (%) | Votes for Clinton (% of population) |
| Median household income (log) | Votes for Trump (% of 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.3: Machine-chosen controls

| | LASSO-chosen controls | Number of controls not chosen |
|---|---|---|
| *County-level analysis* | | |
| Women | Democratic Party Vote Share (2014), Republican Party Vote Share (2014), Votes for DP 2014 (% of population), Unemployment Rate (%), Second Decile Population, Ninth Decile Population | 28 |
| 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.4: Alternative specifications for the first-stage

| | Protesters (%) | | | Protesters (thousands) | | | Log protesters |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| LASSO-chosen weather variable | −0.51 | −0.23 | −0.12 | −41.44 | −19.13 | −40.74 | −0.47 |
| | (0.12) | (0.09) | (0.44) | (13.58) | (5.83) | (11.71) | (0.11) |
| Counties | 2,940 | 2,940 | 470 | 2,940 | 2,940 | 470 | 441 |
| F-Statistic | 17.07 | 6.99 | 0.08 | 9.32 | 10.77 | 12.11 | 17.46 |
| 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 | 1.00 | 0.79 | 1.98 | 1.06 | 0.84 | 6.62 | 0.99 |

Note: The unit of analysis is 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. 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.

Table A.5: Temperature shocks in previous years

| | Temperature shock in January 21 of year: | | | | | | |
|---|---|---|---|---|---|---|---|
| | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
| *Dependent variable* | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *Protesters population (%)* | 0.28 | 0.07 | -0.03 | 0.19 | -0.00 | 0.41 | -0.51 |
| | (0.23) | (0.11) | (0.15) | (0.27) | (0.22) | (0.19) | (0.12) |
| *Women's vote shares (%)* | 2.78 | -1.98 | -0.64 | 0.30 | -0.71 | 2.31 | -4.95 |
| | (1.95) | (1.56) | (2.09) | (1.83) | (2.33) | (2.02) | (1.28) |
| *Underrepresented groups vote shares (%)* | 1.24 | 0.27 | 1.80 | 1.65 | 2.23 | 0.44 | -5.30 |
| | (1.49) | (1.55) | (2.43) | (1.24) | (2.31) | (2.03) | (1.30) |
| *Turnout (%)* | -0.09 | -0.05 | -0.36 | 0.69 | 0.09 | 1.41 | -0.81 |
| | (0.73) | (0.42) | (0.42) | (0.34) | (0.38) | (0.46) | (0.27) |

Note: The unit of analysis is a county and each coefficient (s.e.) comes from a separate first-stage or reduced form regression in which we measure the temperature shock in different years. The temperature shock is January 21'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. Column 7 corresponds to the first-stage and reduced forms in the main analysis of the paper. Standard errors in parentheses are clustered at the state level.

Table A.6: Robustness of results to spatial correlation

| | Vote shares (%) | | |
|---|---|---|---|
| | Women | All underrepresented groups | Turnout (%) |
| | (1) | (2) | (3) |
| *Panel A – Distance cutoff: 100 kms* | | | |
| Protesters (%) | 9.73 | 12.95 | 1.52 |
| | (3.08) | (6.58) | (0.50) |
| *Panel B – Distance cutoff: 50 kms* | | | |
| Protesters (%) | 9.73 | 12.95 | 1.52 |
| | (4.53) | (8.36) | (0.55) |
| Counties | 2,940 | 2,940 | 2,940 |
| Avg. dependent variable | 27.90 | 41.30 | 35.02 |

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. We use distance cutoffs for the spatial kernel of 100kms in Panel A and 50kms in Panel B. The unit of analysis is a county. All regressions are population weighted and include LASSO-chosen controls for each specification. Robust standard errors in parentheses, adjusted for spatial correlation.

Table A.7: Robustness of results to human-selected controls

| | Vote shares (%) | | |
|---|---|---|---|
| | Women | All underrepresented groups | Turnout (%) |
| | (1) | (2) | (3) |
| *Panel A – Reduced Form* | | | |
| LASSO-Chosen weather variable | −3.26 | −5.46 | −0.71 |
| | (1.57) | (1.61) | (0.29) |
| *Panel B – Two-stage least squares* | | | |
| Protesters (%) | 9.87 | 16.52 | 2.13 |
| | (7.10) | (9.79) | (1.57) |
| *Panel C – Ordinary least squares* | | | |
| Protesters (%) | 0.59 | 0.39 | 0.02 |
| | (0.41) | (0.31) | (0.06) |
| Counties | 2,940 | 2,940 | 2,940 |
| Avg. dependent variable | 27.90 | 41.30 | 35.02 |

Note: LASSO-Chosen weather variable is a temperature shock on January 21, 2017. The outcomes are the vote shares obtained by women candidates and candidates from underrepresented group in politics: Women, Hispanic, African-American, Asians/Pacific Islanders or Native Americans, and turnout for the 2018 House of Representatives Election. The unit of analysis is a county. 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.8: 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 Gainesville 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.9: Local reports in counties with *low* temperature shocks

| County ID | Value of the instrument | Protesters (%) | Local report | Local newspaper |
|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (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.10: Robustness of 2SLS results to excluding outliers based on their DFBETA

| | Vote shares (%) | | |
|---|---|---|---|
| | Women | All Underrepresented groups | Turnout (%) |
| | (1) | (2) | (3) |
| Protesters (%) | 7.85 | 10.11 | 1.26 |
| | (2.43) | (4.87) | (0.54) |
| Counties | 2,751 | 2,751 | 2,751 |
| Avg. dependent variable | 27.90 | 41.30 | 35.02 |

Note: This table shows the effect of protests, instrumented with a LASSO-chosen instrument, on the Electoral Outcomes when excluding observations based on their DFBETA. Following the standard approach, we exclude all observations for which $|DFBETA_i| < \frac{2}{\sqrt(N)}$ where N is the number of observations. The unit of analysis is a county. All regressions are population weighted and include LASSO-chosen controls for each specification. Robust standard errors in parentheses, clustered at the state level.