Pro-climate voting in response to local flooding

Online appendix: Supplementary material

Søren Damsbo-Svendsen

Contents

A Research design overview

Table S1: Data, research design, and key variables.

	Study I – party choice	**Study II – candidate choice**
Data	(a) Administrative election results at district-party level, (b) storm surge insurance data	(a) Administrative election results at district-candidate level, (b) storm surge insurance data, (c) candidate survey for polling advice application
Outcome	Party vote share (0-100%)	Candidate vote share (0-100%)
Elections	Danish parliamentary elections: 2011 & 2015	Danish municipal elections: 2013 & 2017
Geounit	822 coastal polling districts (1,386 total)	795 coastal polling districts (1,365 total)*
Unit of analysis	Election-district	Election-district-candidate
Flooding treatment	30-300 worst hit districts (70 highlighted)	30-300 worst hit districts (70 highlighted)
Design	Difference-in-differences (DID)	Triple-differences (DIDID)

Note: *The number of polling districts is slightly lower in Study 2 due to survey non-response and incomplete data merges.

B Data and design details

B.1 Details on the flooding treatment used for Study 1 and 2

1. On a dataset with one row per polling district, I fit a factor model on the three district flooding measures, total cases, rehousing cases, and economic damages, and specify that only one factor should be output.
2. From this factor model, I extract observation-level scores, construed as a latent flooding intensity variable, and add them to the polling district dataset.[1]
3. For each latent intensity score, I calculate how many districts that have a flooding intensity equal to or higher than that threshold. This reflects the number of worst-hit districts that would be in the treatment group, if that particular flooding intensity was used as the threshold.
4. This N, which grows with falling flooding intensity and falls with growing flooding intensity, is used to operationalize the binary treatment variables.[2]
5. I use the 28 thresholds between 300 to 30 in -10 increments in the analysis. Figure S1 shows the level of the three concrete flooding measures for each treatment-control allocation.
6. In certain places, I focus on the N = 70 threshold to simplify the analyses.

[1]Correlations between the latent treatment intensity variable and the original flooding measures are .953, .976, and .999, respectively.

[2]Binary treatment variables are preferred because of the methodological issues with continuous treatments in DID analysis (Callaway et al., 2021).

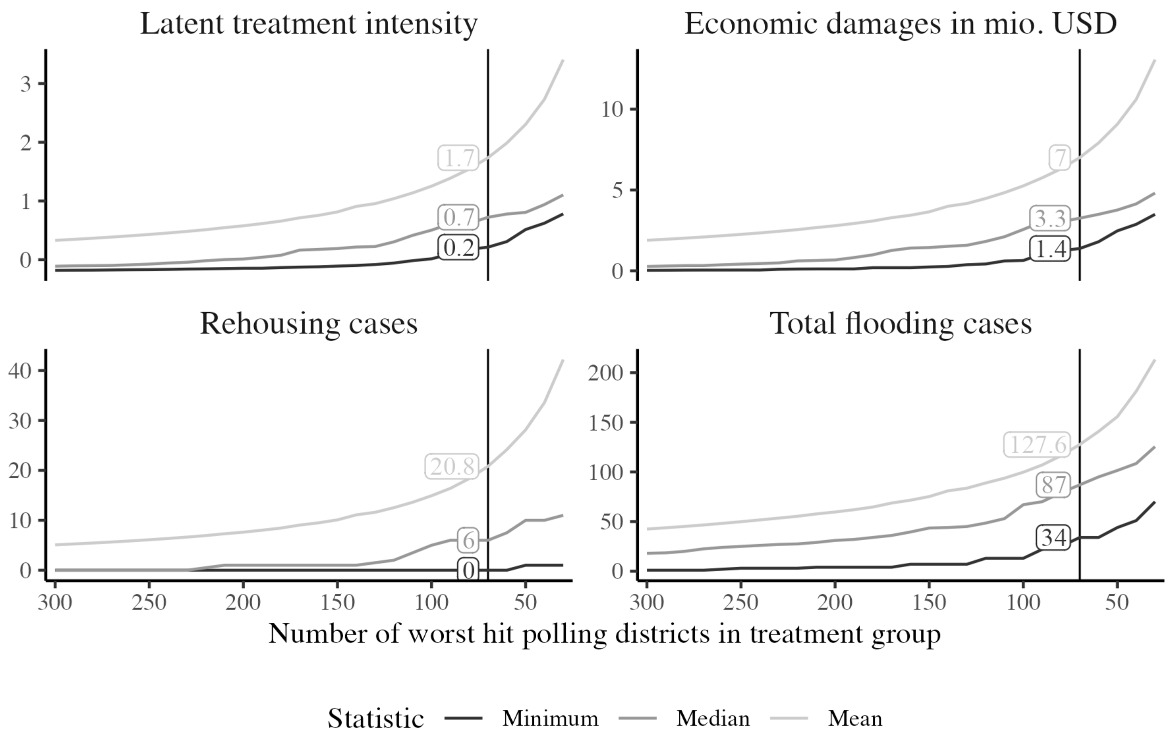

Figure S1: Flooding statistics for each definition of the treatment and control group. The vertical line marks the exemplary treatment at N = 70.

B.2 Details on the district-level vote tallies used for Study 1 and 2

District-level vote tallies for parties and individual election candidates are collected from the Danish Election Database at https://valgdatabase.dst.dk/?lang=en (Thomsen et al., 2022). The issue of adjusted district boundaries, which often impedes researchers' ability to reliably compare election outcomes over time, is solved in the Danish Election Database, which algorithmically harmonizes the election results using the 2015 national election as the baseline.

B.3 Details on the candidate dataset used for Study 2

Election results

1. I collect a dataset with candidate-level votes—available via the Danish Election Database (Thomsen et al., 2022) at https://valgdatabase.dst.dk. The cleaned dataset has one row per election-candidate-polling station for each of the municipal elections of 2013 and 2017: 310,000 rows and 18,640 candidates (9,084 in 2013 and 9,556 in 2017). The number of rows relative to candidates indicates that votes can be cast for a candidate at 16.6 different polling stations, on average (varies between municipalities).
2. Based on municipality and polling station *name*, I attach polling station IDs to be used later for merging with flooding cases. This step is straightforward in 298,207 cases. By matching the remaining districts on approximate polling station names within municipalities, I bring the number of successful matches up to 300,195 out of 310,000. 95 districts are effectively excluded (56 in 2013 and 39 in 2017, equivalent to a total of 9,805 rows).
3. I clean up candidate names by removing special characters, spaces, hyphens, etc.: "René Hansen-Jensen" becomes "ReneHansenJensen" and "Søren Müller" becomes "SorenMuller". I also create "synthetic names" composed of (a) the first 5 and last 4 letters in the clean name or (b) the first 10 letters of the clean name: "René Hansen-Jensen" becomes "ReneHnsen" or "ReneHansen", and "Søren Müller" becomes "Soren-ller" or "SorenMulle".
4. I merge the two datasets, one per election. I do this on the reduced dataset with the 18,640 candidates—one row per election-candidate (not multiplied by the number of polling districts). A certain share of candidates are the same person running in both elections (we do not yet know who). Note that the match of candidates between elections (i.e., 'rerunners') can be performed only from candidate names in combination with municipality and party. It can never be perfect, and it requires assumptions about how registered names, party affiliations, and municipalities (residency) change.
    - From registry data, I know that 4,437 (48.8%) of those running in 2013 were running again in 2017. Ideally, we are therefore able to match most of these 48.8% of 2013 candidates to 2017 candidates.

- 7,082 candidates match perfectly on municipality-party-name. These 3,541 unique people constitute 79.8% of the candidates that should find a match. This is the target.
- Additional 1,348 candidates match on (a) municipality-party and synthetic name, (b) municipality and name (not party), or (c) municipality and synthetic name. (a) Captures the 736 who are likely the same people registered with different spelling or variations of their name (e.g., leaving out one last name); (b) captures 552 party switchers or candidates from small local parties with inconsistent party names (only those matches with just one occurrence per election are included to be safe). From registry data, I know that 780 of the candidates running in both elections change party, but stay in the same municipality. (c) captures a similar group (60 candidates) as (b), but allows the candidate name to vary slightly (I exclude matches with the top 20 most common last names to be safe).
- Finally, 130 candidates are matched based on having a unique full name with only one national occurrence per election. These candidates are thus the only ones allowed to move both municipality and party (they do not have to). From registry data, I know that 966 candidates either move or change party, but only 28 do both.

5. With this procedure, 4,279 candidates running in 2013 and 2017 (96.4%) are successfully matched, and 45.9% of the candidate dataset thus consists of matching pairs (8,558 rows out of 18,640). Matching can never be perfect with the available information, but I have ensured a low rate of false positive matching—assumedly below 1%—and maximized the identification of true positive matches—assumedly close to 96% (4,279 out of 4,437).

Candidate survey

The candidate survey informs a widely used online polling advice application created by *Altinget* and *DR* (Nielsen, 2021). Polling advice applications are highly influential in Danish local elections (Hansen, 2017).

1. I collect two survey datasets about candidates:
   - (a) background information and self-reported, open-ended text field presentations of candidates' key priorities in the 2017 election for 5,863 candidates
   - (b) self-reported, open-ended text field presentations of candidates' key priorities in the 2013 election for 6,089 candidates
   - both datasets contain municipality ID, party name, and candidate name, which I use as matching keys
2. I clean up municipality IDs, party names, candidate names in the same way as in the vote data above, create synthetic names, and clean up the strings with self-reported key priorities (more details below) and some additional candidate information
3. I merge the two datasets onto the candidate-level voting data (within elections). 11,338

(94.9%) match perfectly on election-municipality-party and candidate name, and additional 436 (3.9%) match on election-municipality-party and synthetic name. This brings the total number of matches to 11,774, which is 98.5% of the survey dataset and 63.2% of all candidates in the complete candidate-level voting data (11,774 out of 18,640).

The effective response rate is 63.2% and reflects the share of candidates who gave valid survey responses and also successfully and unambiguously match with vote shares and storm surge cases. It is a high but far from perfect response rate. Therefore, it is important to consider systematic non-participation. Specifically, my strategy is to compare the number of candidates in the complete dataset of candidate votes and the merged dataset of votes and survey responses by (a) municipality and (b) party, which correlate with many factors that may explain non-participation.

Participation is low in very small municipalities with few candidates (down to 29%) and highest in Copenhagen (76%), which is the largest municipality. Participation is higher than 50% in 88 out of 98 municipalities, mean participation is 60%, and median participation is 61%. Most importantly, there is no clear geographical pattern, except that participation rates tend to be higher in very large municipalities and smaller in very small municipalities. The distribution between western and eastern Denmark, for example, is roughly even within the 20 municipalities with highest and lowest participation rates.

Between the nine national parties, survey participation ranges from 46.7% in New Right (radical right party) to 72.5% in the Social Democratic Party (center-left party). Average participation is 61.8%, median participation is 64.3%. There seems to be some tendency for candidates from parties from the fringes of the party system to have lower participation (the three parties with lowest participation are two radical right parties and one socialist party). However, this pattern, which could be an issue, is not consistent: for instance, number two and three on the list are challenger parties on the left and right (a green party, The Alternative, and a liberal party, Liberal Alliance) and number of six out of nine on the list is the traditional governing center-right party, the Liberal Party. The party-level change in participation rates between 2013 and 2017 is also small (except for a 20% drop in participation rates in the Liberal Party). In sum, the variation in participation between parties is potentially problematic, but fortunately there is no systematic pattern, e.g., based on the left-right dimension, the core-fringe dimension, the new-old dimension, etc.

Operationalizing pro-climate candidates

1. Candidates' statements and descriptions of their key priorities are first cleaned by converting them to lower-case, removing all special characters, punctuation marks, etc., and, if applicable, collapsing them into a single string.
2. The statements vary in length. Text lengths reflects a decision by the candidate and is, thus, a valid feature of the data. Nevertheless, it is useful to slightly standardize

the strings. I do this by extracting the first 1,257 characters (approx. 200 words) of every statement corresponding to the 95 percentile.

3. I create a list of non-partisan, climate-related keywords in an iterative manner:
   - (a) I start from the fairly narrow set of keywords used in existing research to identify climate news coverage (Damsbo-Svendsen, 2022) and convert it into a regular expression search string: “\bklima|global[e]? opvarmning|co2|drivhuseffekt|drivhusgas|climate| global warming”
   - (b) I add keywords related specifically to storm surge and flooding, even if they are not among the most frequent terms: “kystsikring|stormflod|oversvømme”
   - (c) I iteratively look up the most commonly used bigrams in the identified strings and add the most relevant among those that are used frequently: “fossile brænd|vedvarende energi|ren energi|alternativ[e]? energi|grøn[n]?[e]? energi|grøn[n]?[e]? løsning|grøn[n]?[e]? omstilling|grøn[n]?[e]? fremtid|grøn[n]?[e]? løsning|grøn[n]?[e]? kommun|grøn[n]?[e]? job|grøn[n]?[e]? arbejdsplads|grøn[n]?[e]? virksomhed|grøn[n]?[e]?[r]?[e]? kommun|kommunen[s]? grøn|bæredygtighed”
   - I analyze the keywords in context (KWIC) to stay clear of words related to climate skepticism or opposition to the green transition; for instance, *wind turbines* is clearly a climate keyword, but its mention mostly reflects local opposition
4. Finally, I identify candidates who match the search string and record the match in a dummy variable. A pro-climate candidate is thus one with a positive match with any of the climate keywords.

Table S2: Climate change keyword frequency.

Keyword	Hits
\bklima	874
bæredygtighed	344
grøn[n]?[e]? omstilling	285
co2	245
vedvarende energi	206
grøn[n]?[e]?[r]?[e]? kommun	147
grøn[n]?[e]? energi	144
grøn[n]?[e]? kommun	126
alternativ[e]? energi	50
fossil	48
oversvømme	34
grøn[n]?[e]? løsning	31
grøn[n]?[e]? løsning	31
grøn[n]?[e]? job	29
grøn[n]?[e]? arbejdsplads	21
kommunen[s]? grøn	20
kystsikring	14
global[e]? opvarmning	13
grøn[n]?[e]? virksomhed	12
grøn[n]?[e]? fremtid	11
c02	10
ren[e]? energi	10
stormflod	9
drivhusgas	5
drivhuseffekt	2
climate	0
global warming	0

Note: Keywords are presented as regular expressions. Frequencies counted across all candidates' open-ended descriptions of key priorities: 11,952 texts with approximately 1,559,300 words. In comparison, there are 2,362 matches with the keyword 'welfare', which is used to identify pro-welfare candidates.

Compiling the candidate dataset

1. I start with the merged dataset described above: 11,774 candidates appearing in the administrative votes dataset and the survey dataset
2. I reattach the vote tallies from each polling station, which extends the dataset to

194,039 rows (election-candidate-districts)

3. I merge the dataset with the district-level flooding treatments used previously in Study 1 (based on polling station IDs). 1,365 districts successfully merge, 78 (5%) do not and are therefore excluded. Importantly, this does not completely exclude any candidates, but only a minor fraction of the within-candidate variation. This reduces the dataset to 187,899 rows.
4. As in Study 1, I restrict the dataset to coastal polling districts, which reduces the final dataset to 118,176 rows (9.748 unique election-candidates and 7.671 unique persons) with variables including candidate ID, election, municipality (where they reside and run for office), polling district, party (unaffiliated candidates and tiny local lists are combined in *other parties*), candidate vote share, candidate ballot position, whether or not they are a pro-climate candidate, pro-welfare candidate, or incumbent (in 2017), and information about storm surge cases (binary treatment variables and flooding intensity measures).

C Supplementary analysis

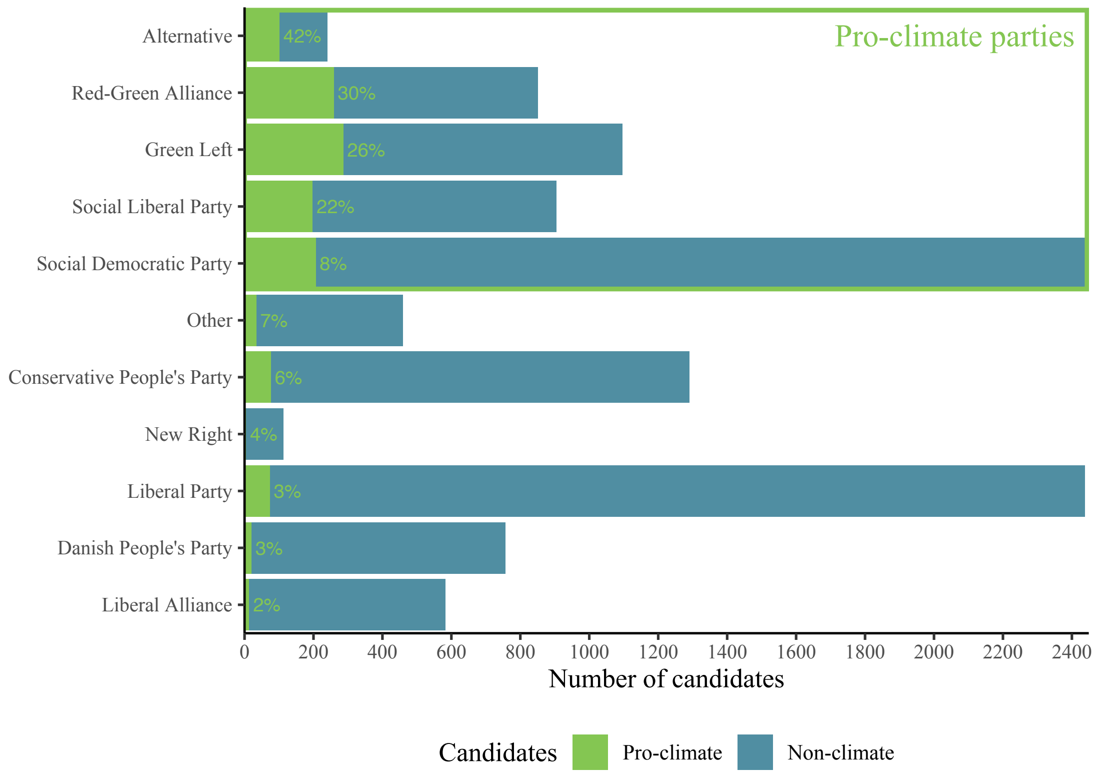

Figure S2: Proportion of pro-climate candidates in the 2013 and 2017 municipal elections by party (Other contains various small regional and local parties). Sorted by pro-climate percentage. N = 11,175 candidates.

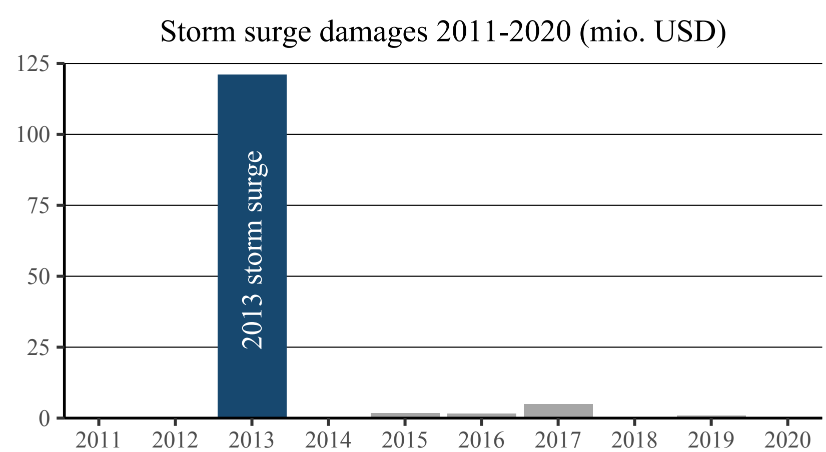

Figure S3: Storm surge flooding in Denmark: annual economic damage reported under the Danish Natural Hazards Council insurance scheme, 2011-2020.

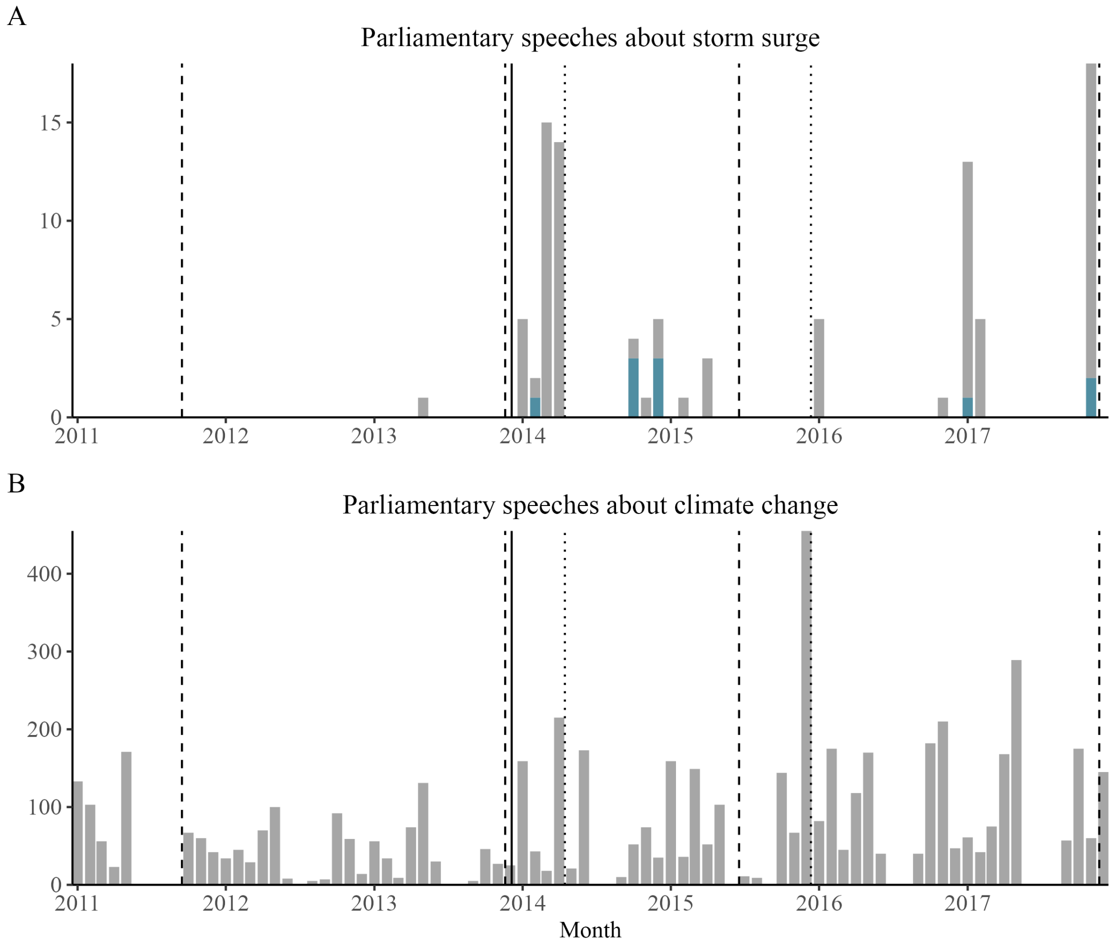

Figure S4: Parliamentary speeches about storm surge and climate change 2011-2017. Panel A: MP mentions storm surge with (blue) or without (grey) explicit reference to the 2013 storm ('Bodil') or the heavily flooded areas ('Roskilde Fjord', 'Jyllinge'). Panel B: MP mentions climate change ('climate', 'global warming', 'CO2', 'greenhouse gas', 'greenhouse effect', 'sea level'). The solid line marks the 2013 storm surge; the dashed lines mark the elections in 2011, 2013, 2015, and 2017; the dotted lines mark the adoption of new Danish storm surge legislation in 2014 and the adoption of the Paris Agreement in 2015. Data from Hjorth (2021).

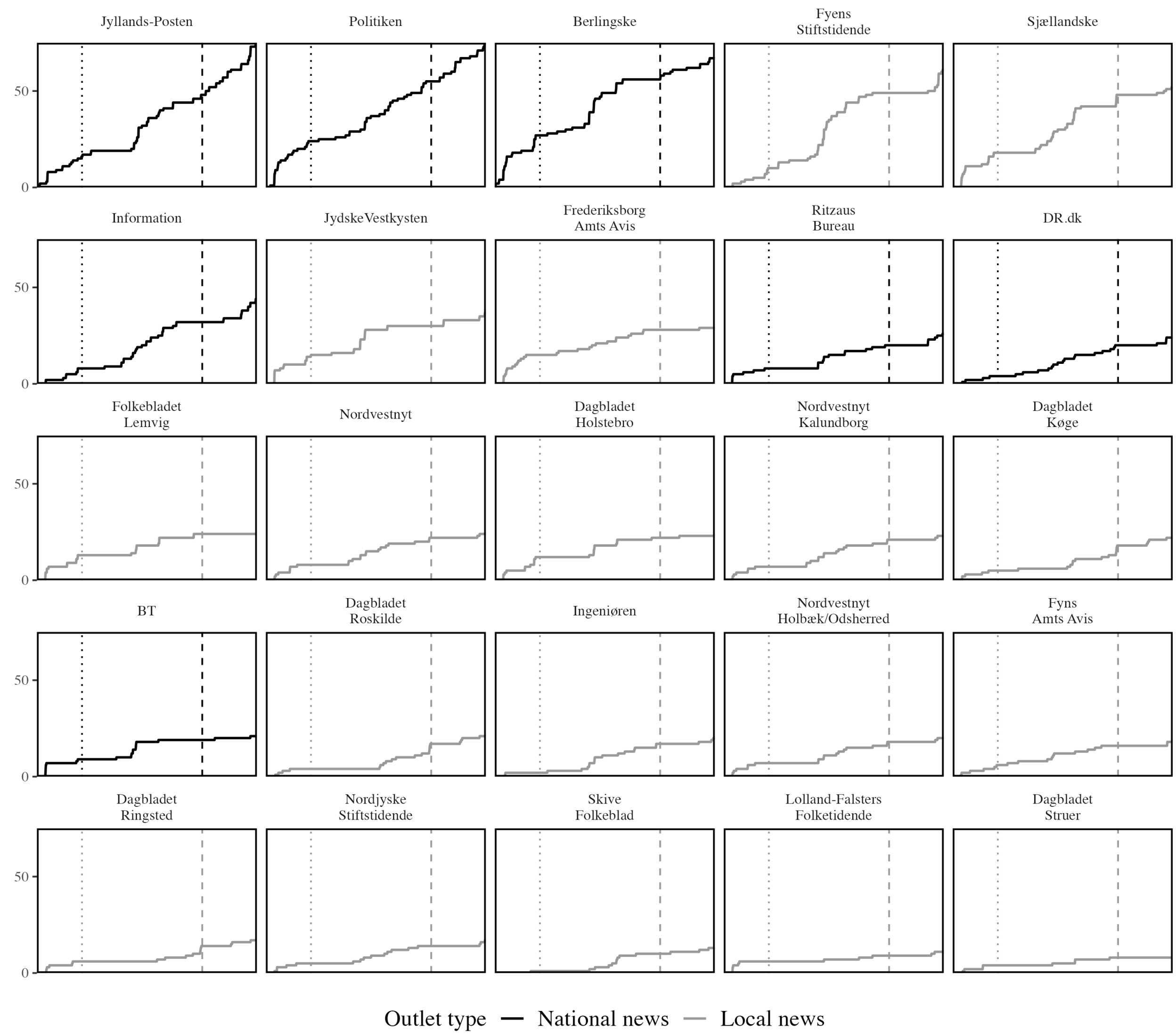

Figure S5: Cumulative number of news articles about the 2013 storm surge in 25 Danish media outlets that covered it most. The unit of measurement is news articles that mention 'storm surge' and the Danish name for the storm ('Bodil'). The time series range from a month before the storm surge (2013-11-05) through the 2015 election year (2015-12-31). The dotted line marks the adoption of new storm surge legislation in the Danish parliament; the dashed line marks the 2015 election. Outlets are sorted by total coverage and colored by type. Data from the media monitoring service, Infomedia.

C.1 Study I – party voting

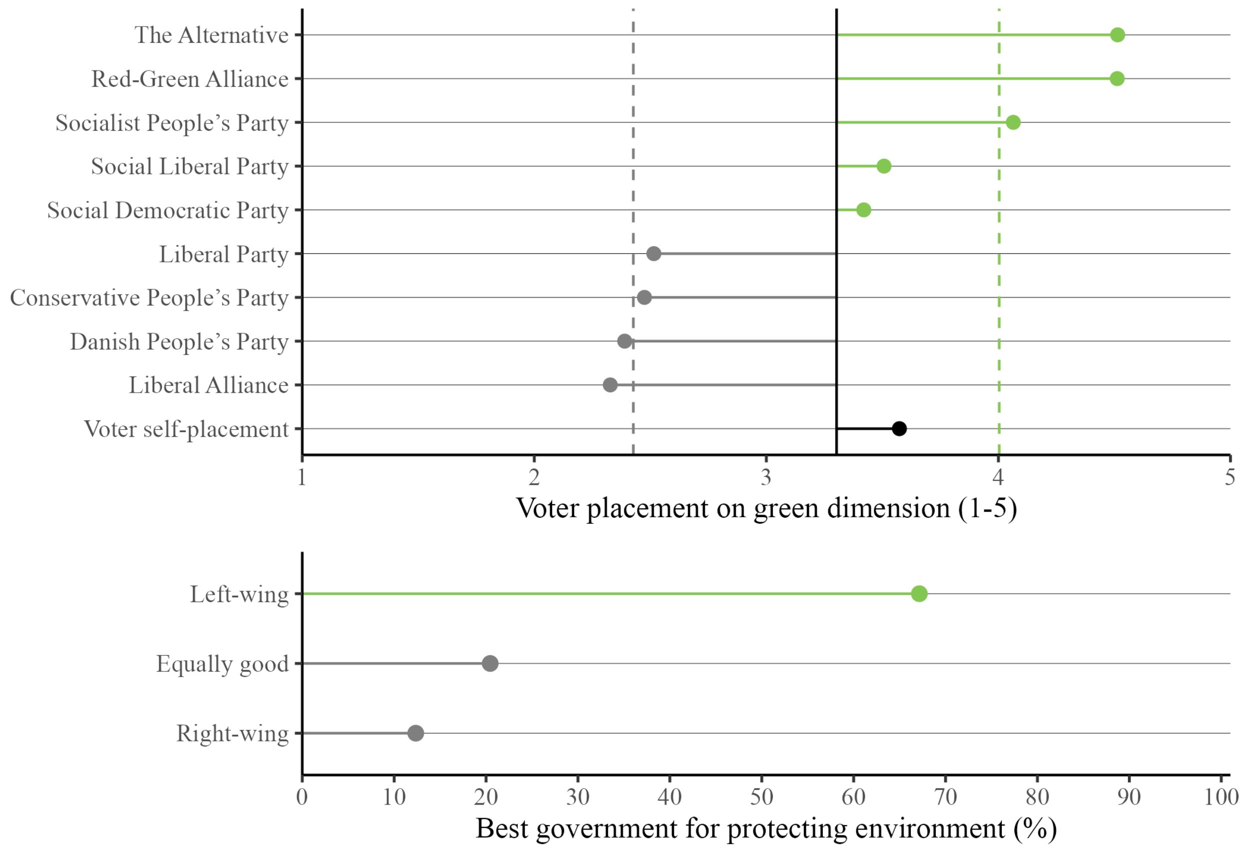

Figure S6: Environmental issue ownership in Denmark. Green points are pro-climate parties, grey points are non-climate parties, and the black point is the average voter self-placement. Incomplete responses on the underlying survey items are excluded. The dashed lines represent group averages, and the solid black line represents the grand mean. Data from the Danish National Election Study 2011 and 2015, pooled. N = 6,225 respondents.

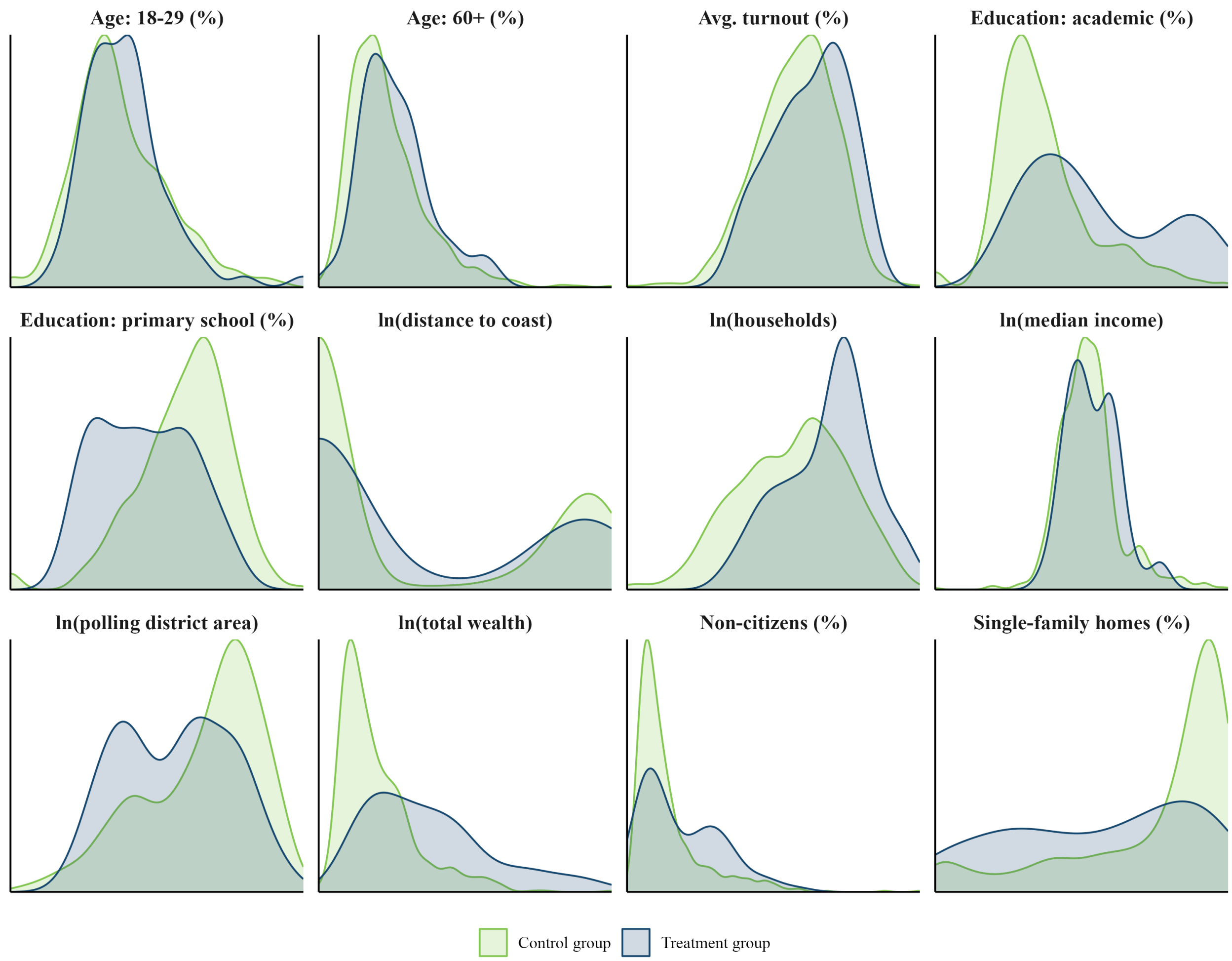

Figure S7: Pre-treatment covariate balance between flooded polling districts (treatment group) and unaffected districts (control group) measured around the 2011 election. Treatment defined as the 70 worst hit districts. The density plots have free scales. N = 805 coastal districts.

Table S3: The effect of local flooding on the pro-climate party vote share: (a) raw DID estimates and (b) matching-based DID estimates.

(a)

	Vote share (%)
	Treatment = 70
Post	-3.54*** (0.12)
Flooding	0.10 (1.31)
Flooding x Post	**3.31*** (0.33)**
Constant	48.54*** (0.36)
R^2	0.031
Standard errors	by: pplace_id
Observations	1614

Note: Main DID effect estimate based on OLS regression. Control group is non-flooded coastal districts. N = 1,614 election-districts. Parliamentary elections of 2015 (Post) and 2011 (ref.). Standard errors clustered at polling district (pplace_id) in parantheses. *p<0.05; **p<0.01; ***p<0.001.

(b)

	Vote share (%)	
		Treatment = 70
95...2	Standard errors	by: district (bootstrapped)
...3	Observations	809

Note: Main DID effect estimate based on inverse probability weighted matching DID using the did package in R. Control group is non-flooded coastal districts. Covariates used for matching include: polling district area, median income, and total wealth, as logged values, and non-citizens, academic education, primary education, single-family homes, and historical average turnout, as percentages. Parliamentary elections of 2015 (Post) and 2011 (ref.). N = 1,614 election-districts across 2011 and 2015 election (809 districts in 2015 and 805 in 2011). Bootstrapped standard errors in parantheses with 95% CIs in square brackets.

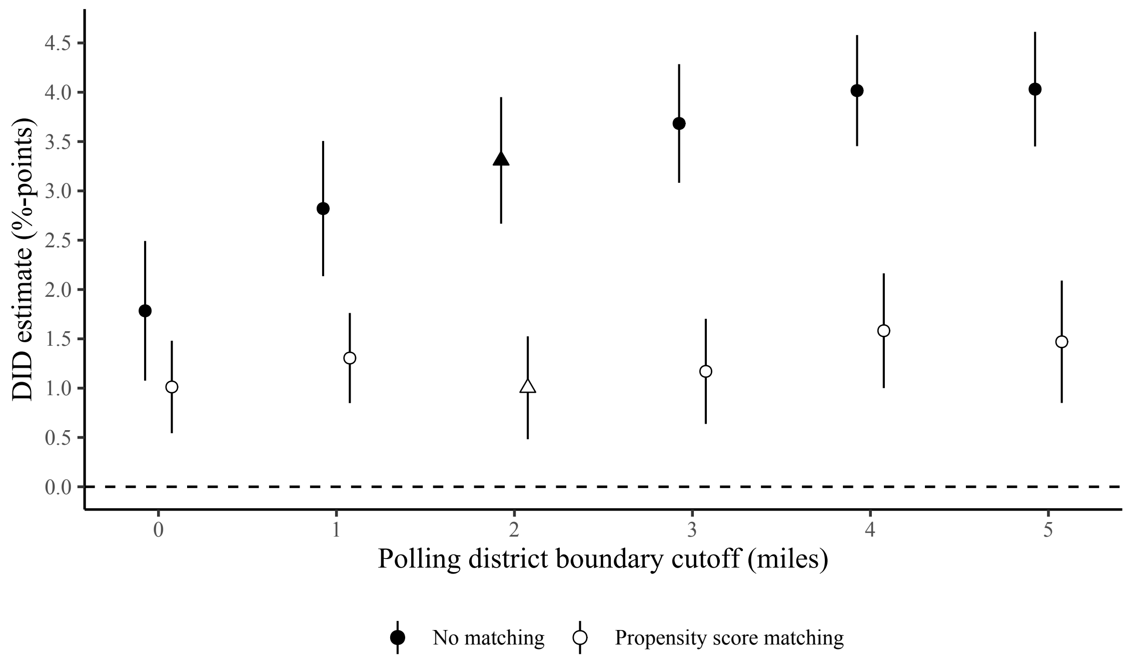

Figure S8: Varying the polling district buffer zone: DID effect estimates of flooding on the pro-climate party vote share. The flooding treatment is defined as the 70 worst hit districts. Six different district boundaries, where 0 implies that only flooding cases within the polling district boundary are included, and 2 (the default) includes cases just outside the district. The control group is non-flooded coastal districts. 95% CIs from standard errors clustered at districts (no matching) or bootstrapped (PS matching). N = 1,614 election-districts.

C.2 Study II – candidate voting

Table S4: Triple differences: Vote shares of local pro-climate candidates versus non-climate candidates after storm surge.

	Candidate vote share (%)
	Treatment = 70
Flooding	0.039*
	(0.015)
Post	-0.031**
	(0.010)
Pro-climate	-0.266***
	(0.032)
Flooding x Post	-0.057*
	(0.024)
Flooding x Pro-climate	-0.147
	(0.085)
Pro-climate x Post	-0.006
	(0.038)
Flooding x Post x Pro-climate	**0.126***
	(0.057)
Municipality FEs	X
Observations	118176

Note: Flooding (dummy) is defined as the 70 worst hit polling districts. Pro-climate candidate (dummy) is defined from mentions of climate-related issues in the candidate survey. Municipal elections of 2017 (Post) and 2013 (ref.). Standard errors clustered at polling district in parantheses. *p<0.05; **p<0.01; ***p<0.001.

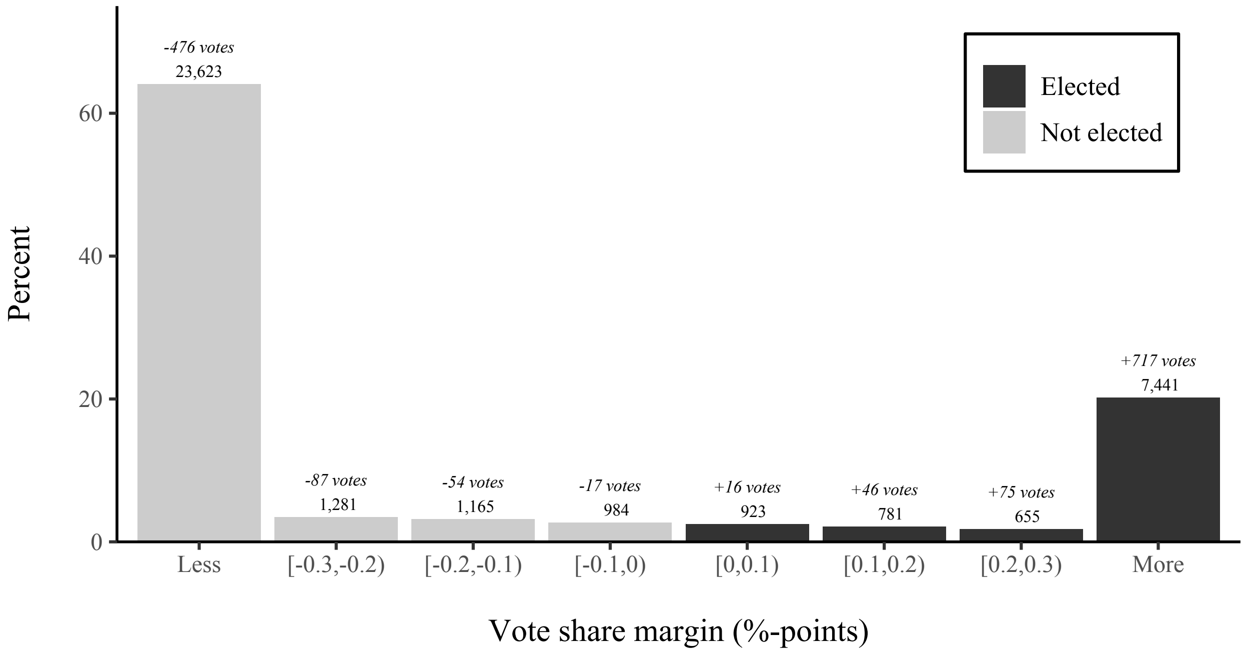

Figure S9: Candidates in the Danish 2009, 2013, 2017, and 2021 local elections: Marginal seat wins and losses. Margin of win or loss on the x-axis. The text above each bar indicates the number of candidates in the group out of a total of 36,853 along with the average win or lose margin within that group.

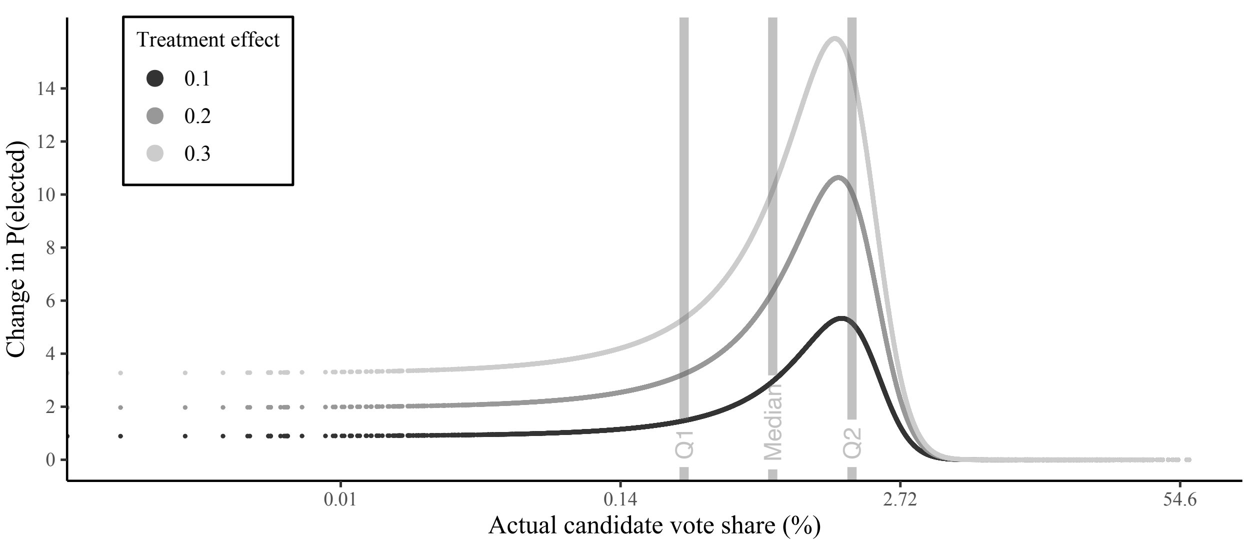

Figure S10: How vote share gains translate into greater election chances. Logit regression of being elected on district-level vote shares. The difference in predicted election chances of every candidate with and without the specified vote share increase. Data is the complete candidate dataset from the 2013 and 2017 local elections. N = 11,672 election-candidates.

Table S5: Effect of flooding on likelihood of being a new candidate rather than running again.

	New candidate	
	Treatment = 70 (minimum)	Treatment = 70 (median)
Constant	0.486***	0.489***
	(0.010)	(0.009)
Flooding	0.021	0.012
	(0.019)	(0.023)
Observations	4851	4851

Note: OLS regression. The data is restricted to the 2017 municipal election, where the outcome can be reliably measured, and aggregated to one row per candidate because there is no variation in being a new candidate between districts. The flooding treatment variable is one if (1) any of the municipality's polling districts are among the 70 worst hit districts or (2) if the median district in the municipality is among the 70 worst hit districts and zero otherwise. Cluster-robust standard errors at municipality level. *p<0.05; **p<0.01; ***p<0.001.

Table S6: Effect of flooding on candidates' ballot positions.

	Ballot position	
	Treatment = 70 (minimum)	Treatment = 70 (median)
Post	-0.002	-0.002
	(0.004)	(0.004)
Pro-climate	0.054**	0.057***
	(0.018)	(0.016)
Flooding x Post	-0.006	-0.011
	(0.007)	(0.008)
Flooding x Pro-climate	-0.022	-0.075
	(0.035)	(0.055)
Pro-climate x Post	-0.019	-0.019
	(0.025)	(0.023)
Flooding x Post x Pro-climate	0.018	0.042
	(0.041)	(0.046)
Municipality fixed effects	X	X
Observations	9748	9748

Note: OLS regression. The data is aggregated to one row per election-candidate because there is no variation in ballot position between districts. The flooding treatment variable is one if (1) any of the municipality's polling districts are among the 70 worst hit districts or (2) if the median district in the municipality is among the 70 worst hit districts and zero otherwise. Municipal elections of 2017 (Post) and 2013 (ref.). Cluster-robust standard errors at municipality level. *p<0.05; **p<0.01; ***p<0.001.

Table S7: Effects of flooding on likelihood of being pro-climate candidate.

	Pro-climate	
	Treatment = 70 (minimum)	Treatment = 70 (median)
Post	0.034***	0.034***
	(0.007)	(0.006)
Flooding x Post	0.013	0.023
	(0.013)	(0.020)
Municipality fixed effects	X	X
Observations	9748	9748

Note: OLS regression. The data is aggregated to one row per election-candidate because there is no variation in being a pro-climate candidate between districts. The flooding treatment variable is one if (1) any of the municipality's polling districts are among the 70 worst hit districts or (2) if the median district in the municipality is among the 70 worst hit districts and zero otherwise. Municipal elections of 2017 (Post) and 2013 (ref.). Cluster-robust standard errors at municipality level. *p<0.05; **p<0.01; ***p<0.001.

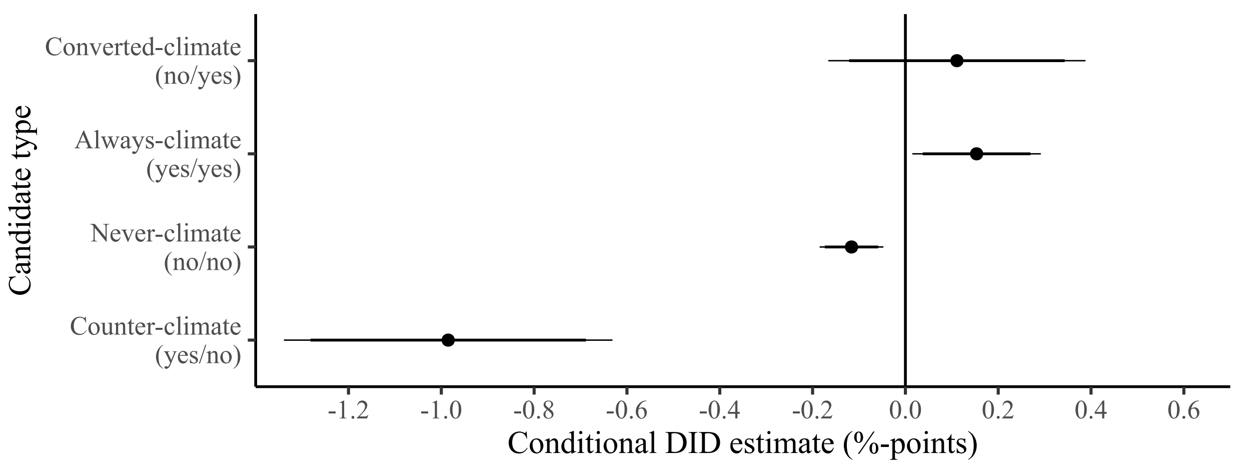

Figure S11: Conditional effect of flooding on candidate vote shares for four types of candidates. Candidate type (categorical) is defined from whether they were pro-climate candidates in 2013 and 2017: always-climate (8 percent), converted-climate (4 percent), never-climate (86 percent), and counter-climate. The data is restricted to the 2,077 unique candidates running in both election. The flooding treatment is defined as the 70 worst hit districts. DID estimates from OLS model with municipality fixed effects. 95% and 90% CIs (standard errors clustered on polling districts). N = 28,085 election-district-candidates.

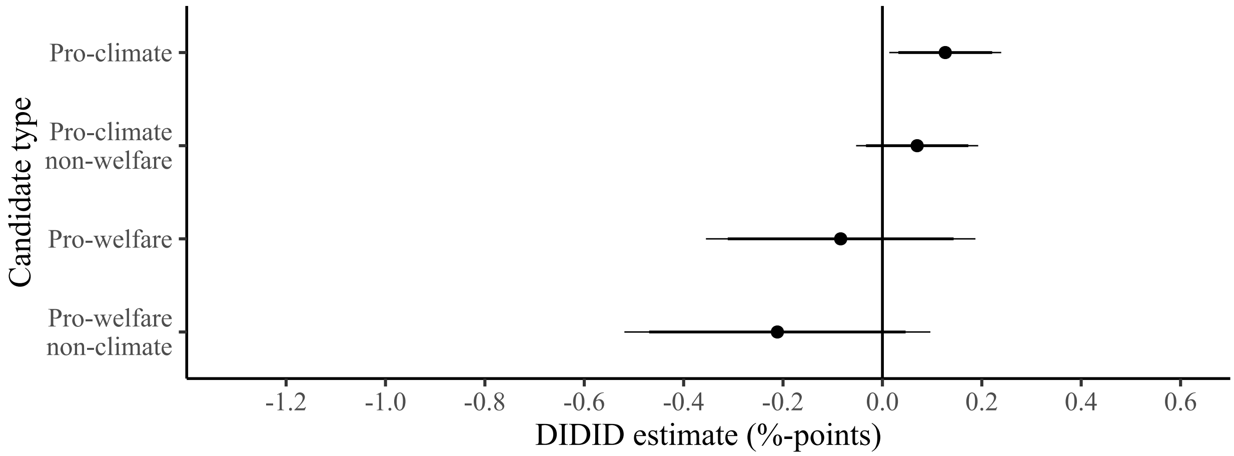

Figure S12: Effects of flooding on pro-climate and pro-welfare candidates' vote shares. The flooding treatment is defined as the 70 worst hit districts. DID estimates from OLS regression model with municipality fixed effects. The figure shows DIDID estimates for pro-climate and pro-welfare candidates and for the part of each category that does not overlap with the other. The pro-welfare dummy variable is coded based on use of the keyword, 'welfare', which is one of the least ambiguous and most frequently used keywords. Approximately 9.3 percent of the sample are only pro-climate, 11.7 percent are only pro-welfare, 1.7 percent are both, and 77.3 percent are neither. Municipal elections of 2017 (Post) and 2013 (ref.). Cluster-robust standard errors at district level. 95% and 90% CIs (standard errors clustered on polling districts). N = 28,085 election-district-candidates.

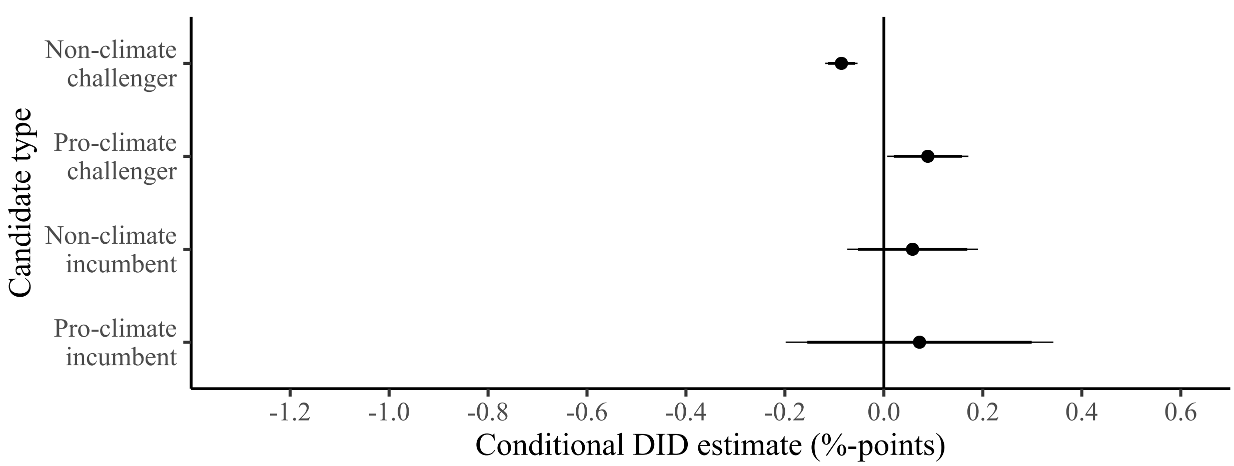

Figure S13: Flooding effect conditional on incumbency and climate communication. Conditional effects of flooding on the vote shares of four candidate types: Non-climate non-incumbents, pro-climate non-incumbents, non-climate incumbents, and pro-climate incumbents. The flooding treatment is defined as the 70 worst hit districts. DID estimates from OLS model with municipality fixed effects. 95% and 90% CIs (standard errors clustered on polling districts). N = 28,085 election-district-candidates.

References

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with a Continuous Treatment. *arXiv:2107.02637*. https://arxiv.org/abs/2107.02637

Damsbo-Svendsen, S. (2022). Mass Media Influence on the Rapid Rise of Climate Change. *International Journal of Public Opinion Research*, *34*(2).

Hansen, K. M. (2017). Partiers og kandidaters forskellige valgkampagner [Parties' and candidates' different election campaigns]. In J. Elklit, C. Elmelund-Præstekjær, & U. Kjær (Eds.), *KV13. Analyser af kommunalvalget 2013* (pp. 61–78). University of Southern Denmark Press.

Hjorth, F. (2021). *Danish Legislative Speech Corpus*. Harvard Dataverse.

Nielsen, J. (2021). Sådan har vi lavet kandidattesten [How we made the candidate test]. *Altinget*.

Thomsen, S. R., Søderberg, J., & Statistics Denmark. (2022). *The Danish Election Database.*