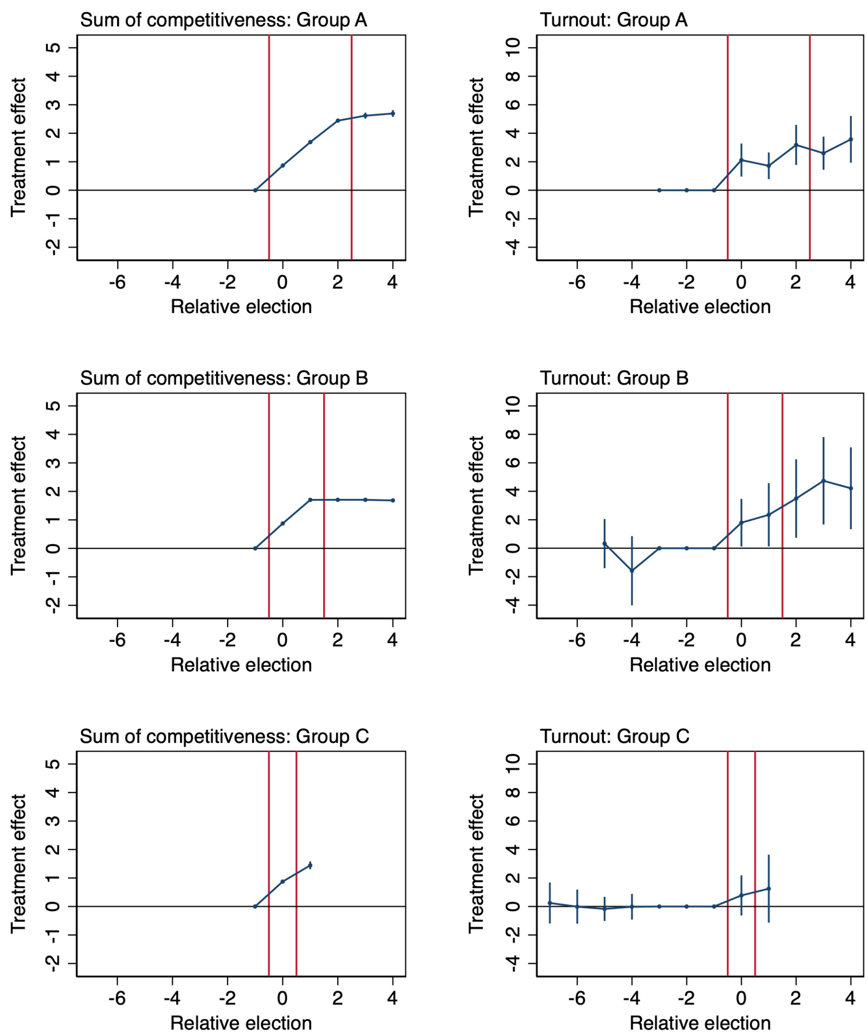
# Online appendix for:

## District competitiveness increases voter turnout: evidence from repeated redistricting in North Carolina

Robert Ainsworth
Emanuel Garcia Munoz
Andres Munoz Gomez

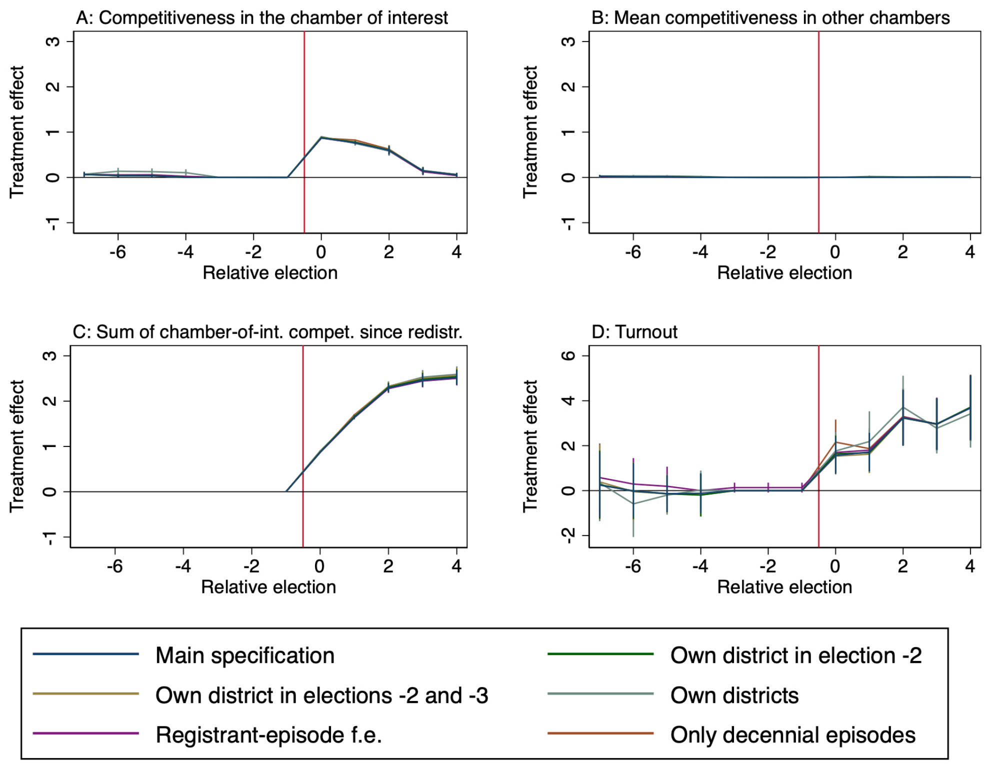
# Additional figures and tables

Figure A1: The effects of assigned competitiveness, $c_{a_i}$, by treatment group



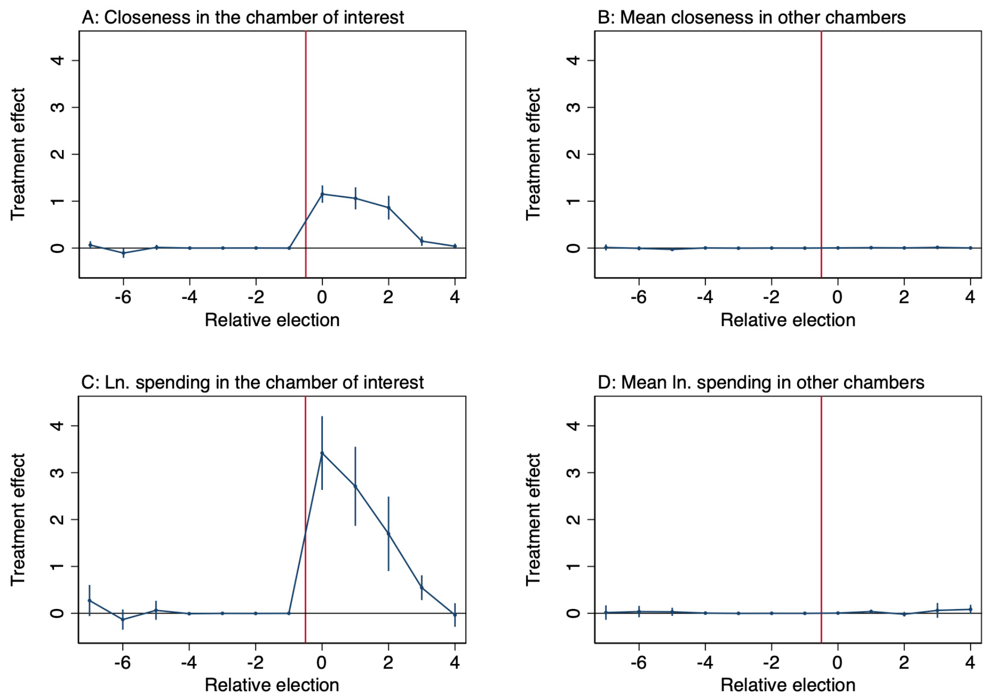

The figure displays event studies that are analogous to those in Panels C and D of Figure 3. The difference is that results are provided separately by "treatment group". A treatment group is defined as a set of redistricting episodes whose districts last for the same length of time. The vertical red lines bound the elections during which a given group's districts are in use. Group A is the episodes with districts that last for three elections. These are the decennial episodes for the NC Senate and NC House. Group B is the episodes with districts that last for two elections: the decennial episode and the first revision for the U.S. House. In Group C, districts last for only one election. These episodes are the first and second revisions for the NC Senate and NC House and the second revision for the U.S. House. "Sum of competitiveness" is $C_{i\tau}$. Turnout is $to_{i\tau}$ and is denominated in percentage points. Results are not provided for $\tau \leq -2$ in the "Sum of competitiveness" panels because $C_{i\tau}$ is 0 in all pre-redistricting elections. Standard errors are clustered by baseline district in the chamber of interest.

Figure A2: Robustness in the effects of assigned competitiveness, $c_{a_i}$




The figure displays event studies that are analogous to those in Figure 3. "Main specification" is the same as that in Figure 3. "Own district in election -2" is similar to the main specification but also matches on a registrant's chamber-of-interest district in $\tau = -2$ (the first election before the baseline). "Own districts in elections -2 and -3" matches on chamber-of-interest districts in both $\tau = -2$ and $\tau = -3$. "Own districts" drops the matching on baseline region and instead matches on a set of a registrant's districts in both pre- and post-redistricting elections in all chambers; see Appendix A3 for details. "Registrant-episode f.e." adds a set of fixed effects for registrant-episode combinations. "Only decennial episodes" creates a balanced panel by using only the decennial redistricting episodes. All other details are as in Figure 3.

Figure A3: The effects of assigned competitiveness, $c_{a_i}$ on experiences in legislative races




The figure presents results that are analogous to those in Panels A and B of Figure 3. However, the outcomes relate to registrants' experiences in legislative races, not the competitiveness of registrants' districts. "Closeness" is race closeness, defined as 1 minus the absolute two-party vote-share margin. "Ln. spending" is the natural log of per-person race spending, measured in 2010 dollars.

Table A1: Robustness to alternative instruments

| | Main | Alternative instruments | | |
|---|---|---|---|---|
| | | (1) | (2) | (3) |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | $1.30^{***}$ | $1.30^{***}$ | $1.30^{***}$ | $1.30^{***}$ |
| | (0.234) | (0.234) | (0.232) | (0.230) |
| Turnout percentage | 58.1 | 58.1 | 58.1 | 58.1 |
| Clusters | 338 | 338 | 338 | 338 |
| Registrants | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 |
| Registrant-episode-elections | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 |

The table presents results for alternative ways of constructing instruments. Specifically, it shows coefficient estimates and standard errors for $\alpha$ from versions of Model (3) that use different first-stage specifications. "Main" is for the main specification, as in the "All" column of Table 1. This specification allows the first-stage coefficient, $\beta$, to vary by $\tau$. Thus, the instruments in this specification are the interaction of $C_{a_i\tau}$ and indicators for $\tau$. The only instrument in Column 1 is $C_{a_i\tau}$; thus, the specification in this column forces $\beta$ to be the same for all $\tau$. The instruments in Column 2 are the interaction of $C_{a_i\tau}$ and indicators for treatment-group-by-$\tau$. Those in Column 3 are the interaction of $C_{a_i\tau}$ and indicators for episode-by-$\tau$. See Appendix A4 for more details on instruments. All other details are the same as in Table 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Robustness to alternative competitiveness measures

| | (1) | (2) | (3) |
|---|---|---|---|
| Sum of standardized competitiveness in a registrant's districts: main measure | $1.30^{***}$ | | |
| | (0.234) | | |
| Sum of standardized competitiveness in a registrant's districts: alternative measure | | $1.13^{***}$ | |
| | | (0.239) | |
| Sum of standardized competitiveness in a registrant's districts: Cook measure | | | $1.38^{***}$ |
| | | | (0.271) |
| Standard deviation of $c_{i\tau}$ | 0.139 | 0.159 | 0.127 |
| Effect of a 1 s.d. increase in $c_{i\tau}$ | 0.181 | 0.180 | 0.176 |
| Turnout percentage | 58.1 | 58.1 | 58.1 |
| Clusters | 338 | 338 | 338 |
| Registrants | 5,203,371 | 5,203,371 | 5,203,371 |
| Registrant-episode-elections | 31,366,989 | 31,366,989 | 31,366,989 |

The table presents results for versions of Model (3) that use different measures of district competitiveness. In Column 1, $C_{i\tau}$ and $C_{a_i\tau}$ are constructed using the main measure, $c_{d,\mathrm{M}}$. Results in this column correspond with those in the "All" column of Table 1. Columns 2 and 3 use, respectively, the alternative measure, $c_{d,\mathrm{A}}$, and the Cook measure, $c_{d,\mathrm{PVI}}$. See Appendix A2 for details on the competitiveness measures. "Standard deviation of $c_{i\tau}$" is the standard deviation of the listed competitiveness measure, as shown in Table A20. "Effect of a 1 s.d. increase in $c_{i\tau}$" is the product of the coefficient estimate and the standard deviation of $c_{i\tau}$. Other details are the same as in Table 1.

## Table A3: Robustness to alternative match-groups

| | Main | Alternative match-groups | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 1.29*** | 1.29*** | 1.30*** | 1.32*** | 1.13*** | 1.12*** | 1.21*** |
| | (0.234) | (0.247) | (0.230) | (0.235) | (0.233) | (0.217) | (0.231) | (0.210) |
| Turnout percentage | 58.1 | 57.5 | 58.2 | 57.5 | 57.4 | 58.9 | 58.1 | 58.3 |
| Clusters | 338 | 338 | 338 | 338 | 338 | 338 | 338 | 335 |
| Registrants | 5,203,371 | 5,820,059 | 5,319,023 | 5,452,937 | 5,547,793 | 4,859,572 | 4,947,697 | 4,768,478 |
| Registrant-episode-elections | 31,366,989 | 38,955,655 | 32,544,782 | 33,577,639 | 34,375,982 | 28,372,471 | 29,255,856 | 27,094,207 |

The table presents results for alternative ways of constructing match-groups. Specifically, it shows coefficient estimates and standard errors for $\alpha$ from versions of Model (3) that define match-groups using different sets of covariates. "Main" is for the main definition, as in the "All" column of Table 1. The remaining columns add or remove a single covariate. Column 1 removes the share college graduates in a registrant's baseline block-group. Column 2 removes a registrant's turnout behavior in $\tau = -3$. Column 3 removes the registrant's gender. Column 4 removes the registrant's baseline party registration. Column 5 adds two groups for population density in the registrant's baseline census block. Column 6 adds two groups for the value of the registrant's baseline property parcel. Column 7 adds two groups for the median household income in the registrant's baseline block-group. See Appendix A3 for details on match-groups. All other details are the same as in Table 1.

## Table A4: Robustness to matching on additional district variables

| | **Main specification** | **Own district in election $\tau$ = -2** | **Own district in elections -2 and -3** | **Own districts** |
|---|---|---|---|---|
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 1.30*** | 1.27*** | 1.31*** |
| | (0.234) | (0.227) | (0.227) | (0.240) |
| Turnout percentage | 58.1 | 57.9 | 57.5 | 57.6 |
| Clusters | 338 | 338 | 338 | 366 |
| Registrants | 5,203,371 | 5,052,608 | 4,874,586 | 5,035,576 |
| Registrant-episode-elections | 31,366,989 | 30,146,952 | 28,504,402 | 31,434,795 |

The table presents results for versions of Model (3) that match on additional district variables. "Main specification" is the same as in the "All" column of Table 1. "Own district in election $\tau = -2$" modifies the main specification by also matching on a registrant's chamber-of-interest district in $\tau = -2$. "Own districts in elections -2 and -3" matches on chamber-of-interest districts in both $\tau = -2$ and $\tau = -3$. "Own districts" doesn't match on baseline region and instead matches on a set of a registrant's districts in both pre- and post-redistricting elections in all chambers; see Appendix A3 for details. Other details are the same as in Table 1.

Table A5: Robustness in the effects of district competitiveness on consistent turnout

| | Turnout type | | |
|---|---|---|---|
| | Any | Consistent | Inconsistent |
| *Panel A: ≥ 1 future election* | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.28*** | 1.25*** | 0.023 |
| | (0.230) | (0.243) | (0.114) |
| Turnout percentage | 55.3 | 42.1 | 13.2 |
| Clusters | 255 | 255 | 255 |
| Registrants | 4,486,052 | 4,486,052 | 4,486,052 |
| Registrant-episode-elections | 22,597,415 | 22,597,415 | 22,597,415 |
| *Panel B: ≥ 2 future elections* | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.38*** | 1.31*** | 0.067 |
| | (0.260) | (0.265) | (0.183) |
| Turnout percentage | 57.0 | 39.1 | 17.9 |
| Clusters | 163 | 163 | 163 |
| Registrants | 3,928,260 | 3,928,260 | 3,928,260 |
| Registrant-episode-elections | 15,713,852 | 15,713,852 | 15,713,852 |
| *Panel C: ≥ 3 future elections* | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.39*** | 1.47*** | -0.082 |
| | (0.288) | (0.334) | (0.190) |
| Turnout percentage | 54.5 | 35.3 | 19.2 |
| Clusters | 151 | 151 | 151 |
| Registrants | 3,839,532 | 3,839,532 | 3,839,532 |
| Registrant-episode-elections | 10,271,220 | 10,271,220 | 10,271,220 |

The table is similar to Table 3 but restricts the sample to episode-election combinations for which there is data on a given number of later elections. Panel A is for combinations with data on at least one future election. Panel B (C) is for combinations with data on at least two (three) future elections.

Table A6: Heterogeneity in the effects of district competitiveness by registrant characteristics

| | All | Age | | Education | | Gender | |
|---|---|---|---|---|---|---|---|
| | | $\leq 35$ | $\geq 36$ | Low | High | Male | Not male |
| Sum of competitiveness in | 1.30*** | 1.70*** | 1.13*** | 1.20*** | 1.50*** | 1.42*** | 1.21*** |
| a registrant's districts, $C_{i\tau}$ | (0.234) | (0.441) | (0.198) | (0.292) | (0.380) | (0.254) | (0.232) |
| Turnout percentage | 58.1 | 40.6 | 65.8 | 56.1 | 62.1 | 57.0 | 58.9 |
| Clusters | 338 | 337 | 338 | 305 | 217 | 338 | 338 |
| Registrants | 5,203,371 | 1,837,032 | 3,567,465 | 3,640,702 | 1,808,555 | 2,316,641 | 2,888,676 |
| Registrant-episode-elections | 31,366,989 | 9,600,517 | 21,766,472 | 20,994,874 | 10,372,115 | 13,842,727 | 17,524,262 |

The table presents heterogeneity in the effects of district competitiveness by various registrant characteristics. "All" is the main version of Model (3), as in the "All" column of Table 1. The remaining columns show results for versions of Model (3) that are fit using only registrants with the specified traits. "Age" is measured in the baseline election. "Education" is low (high) if the share college graduates in the baseline block-group is less than or equal to (greater than) 0.4. Other details are the same as in Table 1.

Table A7: Heterogeneity by age and education

| | All | Low education | | High education | |
|---|---|---|---|---|---|
| | | $\leq 35$ | $\geq 36$ | $\leq 35$ | $\geq 36$ |
| Sum of competitiveness in | 1.30*** | 1.28*** | 1.16*** | 2.40*** | 1.06*** |
| a registrant's districts, $C_{i\tau}$ | (0.234) | (0.482) | (0.268) | (0.765) | (0.278) |
| Turnout percentage | 58.1 | 38.7 | 63.7 | 44.4 | 69.9 |
| Clusters | 338 | 302 | 305 | 215 | 216 |
| Registrants | 5,203,371 | 1,256,184 | 2,498,381 | 632,437 | 1,231,764 |
| Registrant-episode-elections | 31,366,989 | 6,420,754 | 14,574,120 | 3,179,763 | 7,192,352 |

The table shows how the effects of district competitiveness vary for subsets of registrants defined by the interaction of age and education. Results are for versions of Model (3) that are fit using only registrants of the specified type. See Table A6 for details on the definitions of the age and education variables. See Table 1 for all other details.

Table A8: Heterogeneity by election type

| | All | Election type | |
|---|---|---|---|
| | | Midterm | Presidential |
| Sum of competitiveness in | 1.30*** | 1.12*** | 1.42*** |
| a registrant's districts, $C_{i\tau}$ | (0.234) | (0.217) | (0.272) |
| Turnout percentage | 58.1 | 48.1 | 64.2 |
| Clusters | 338 | 255 | 338 |
| Registrants | 5,203,371 | 4,486,052 | 5,203,371 |
| Registrant-episode-elections | 31,366,989 | 12,019,173 | 19,347,816 |

The table show how the effects of district competitiveness vary for midterm v. presidential elections. Results are for versions of Model (3) that are fit using only elections of the specified type. Other details are the same as in Table 1.

Table A9: Effects for registrants assigned to districts in which their incumbent does not run

| | **All** | **Own incumbent not running** |
|---|---|---|
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 1.50*** |
| | (0.234) | (0.329) |
| Turnout percentage | 58.1 | 58.2 |
| Clusters | 338 | 337 |
| Registrants | 5,203,371 | 4,028,963 |
| Registrant-episode-elections | 31,366,989 | 20,124,500 |

The table shows how the effects of district competitiveness interrelate with the effects of incumbency. "All" presents results for the main version of Model (3), as in the "All" column of Table 1. "Own incumbent not running" restricts the sample to registrants who are assigned to districts in which their pre-redistricting incumbent is not one of the candidates in the first election after redistricting. For these registrants, the effect of district competitiveness is not distorted by attachment to a given incumbent. Note that we cannot run Model (3) just for registrants who are assigned to districts in which their pre-redistricting incumbent does run. This is because registrants in the same match-group share the same pre-redistricting incumbent, and thus there can only be one assigned district per match-group in which the group's incumbent runs.

Table A10: Effects calculated using differences in assigned districts for multiple chambers: testing interactions

| | **Weighted sum** | | | **Sums by chamber** | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| Weighted sum of competitiveness in a registrant's districts: all chambers | 1.01*** | 1.04*** | 1.08*** | | | |
| | (0.19) | (0.21) | (0.24) | | | |
| Sum of competitiveness in a registrant's districts: U.S. House | | | | 2.57*** | 2.56*** | 2.66*** |
| | | | | (0.77) | (0.74) | (0.78) |
| Sum of competitiveness in a registrant's districts: state chambers | | | | 1.11*** | 1.16*** | 1.21*** |
| | | | | (0.23) | (0.27) | (0.31) |
| Sum of the interaction of competitiveness in the U.S. House & NC Senate | | 7.57* | 6.61 | | 7.73* | 6.87 |
| | | (4.17) | (4.52) | | (4.18) | (4.51) |
| Sum of the interaction of competitiveness in the U.S. House & NC House | | 2.95 | 3.00 | | 3.22 | 3.28 |
| | | (3.34) | (3.31) | | (3.38) | (3.35) |
| Sum of the interaction of competitiveness in the NC Senate & NC House | | 0.68 | 0.74 | | 0.42 | 0.47 |
| | | (2.05) | (2.06) | | (1.99) | (2.00) |
| Sum of the interaction of competitiveness in all chambers | | | -11.6 | | | -10.4 |
| | | | (20.1) | | | (20.1) |
| Turnout percentage | 58.0 | 58.0 | 58.0 | 58.0 | 58.0 | 58.0 |
| F-stat. for joint significance | - | 2.75 | 2.12 | - | 2.81 | 2.15 |
| p-value for joint significance | - | 0.042 | 0.077 | - | 0.039 | 0.074 |
| Clusters | 540 | 540 | 540 | 540 | 540 | 540 |
| Registrants | 5,604,366 | 5,604,366 | 5,604,366 | 5,604,366 | 5,604,366 | 5,604,366 |
| Registrant-episode-elections | 27,919,274 | 27,919,274 | 27,919,274 | 27,919,274 | 27,919,274 | 27,919,274 |

The table is similar to Table 5 but adds interaction terms. "Weighted sum" ("Sums by chamber") is for models that correspond with Column 1 (2) of the earlier table. The interaction terms are sums of the products of competitiveness in the listed chambers. They are instrumented using analogous sums of the products of assigned competitiveness. We subtract 0.75 from competitiveness variables before calculating products and sums. This way, the coefficients on the non-interaction terms represent marginal effects for registrants in 62.5-37.5 districts, not 100-0 districts. The "F-stat." and "p-value" rows provide results from an F-test for joint significance of the interaction terms. All other details are the same as in Table 5.

Table A11: Summarizing $\bar{c}_i^j$, the average competitiveness of a registrant's districts in chamber $j$

| | **Mean** | **Std. dev.** | **Percentile** | |
|---|---|---|---|---|
| | | | 10th | 90th |
| *Panel A: U.S. House* | | | | |
| All registrants | 0.81 | 0.06 | 0.72 | 0.89 |
| White-Democrats | 0.82 | 0.05 | 0.73 | 0.89 |
| White-Republicans | 0.82 | 0.05 | 0.77 | 0.89 |
| White-Unaffiliated | 0.82 | 0.05 | 0.76 | 0.89 |
| Racial minorities | 0.79 | 0.06 | 0.70 | 0.87 |
| *Panel B: NC Senate* | | | | |
| All registrants | 0.77 | 0.10 | 0.65 | 0.90 |
| White-Democrats | 0.78 | 0.09 | 0.67 | 0.91 |
| White-Republicans | 0.77 | 0.09 | 0.67 | 0.90 |
| White-Unaffiliated | 0.78 | 0.09 | 0.67 | 0.91 |
| Racial minorities | 0.74 | 0.11 | 0.58 | 0.87 |
| *Panel C: NC House* | | | | |
| All registrants | 0.74 | 0.11 | 0.61 | 0.89 |
| White-Democrats | 0.76 | 0.11 | 0.63 | 0.91 |
| White-Republicans | 0.75 | 0.10 | 0.62 | 0.89 |
| White-Unaffiliated | 0.75 | 0.11 | 0.62 | 0.89 |
| Racial minorities | 0.72 | 0.12 | 0.54 | 0.87 |

The table shows how registrants' exposure to competitiveness differed by legislative chamber. It presents summary statistics for chamber-specific versions of $\bar{c}_i$, $\bar{c}_i^j$. These are:

$$\bar{c}_i^j = \frac{1}{5}\sum_{t=2012}^{2020} c_{it}^j.$$

The values in Panel A are for $\bar{c}_i^{\text{USH}}$, while those in B and C are for $\bar{c}_i^{\text{NCS}}$ and $\bar{c}_i^{\text{NCH}}$. See Figure 5 for more details.

Table A12: Summarizing $\bar{c}_{it}$, the average competitiveness of a registrant's districts in election $t$

| **Election** | **Mean** | **Std. dev.** | **Percentile** | |
|---|---|---|---|---|
| | | | 10th | 90th |
| 2012 | 0.75 | 0.09 | 0.63 | 0.87 |
| 2014 | 0.75 | 0.09 | 0.63 | 0.87 |
| 2016 | 0.77 | 0.08 | 0.67 | 0.88 |
| 2018 | 0.79 | 0.08 | 0.69 | 0.89 |
| 2020 | 0.81 | 0.09 | 0.68 | 0.91 |

The table shows how registrants' exposure to competitiveness differed by election. It presents summary statistics for election-specific versions of $\bar{c}_i$, $\bar{c}_{it}$. These are:

$$\bar{c}_{it} = \frac{1}{3}\sum_{j\in\{\text{USH},\text{NCS},\text{NCH}\}} c_{it}^j.$$

See Figure 5 for more details.

## Table A13: The coefficient estimates used in the simulation

| | All | White | | | Minority | Low education | | High education | |
|---|---|---|---|---|---|---|---|---|---|
| | | Dem. | Rep. | Unaffil. | | $\leq 35$ | $\geq 36$ | $\leq 35$ | $\geq 36$ |
| *Panel A: U.S. House* | | | | | | | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 2.03*** | 3.11*** | 2.40*** | 2.55*** | 0.633 | 0.915 | 2.09*** | 4.53* | 2.43** |
| | (0.738) | (0.632) | (0.616) | (0.845) | (2.31) | (1.74) | (0.727) | (2.36) | (1.10) |
| Turnout percentage | 58.3 | 62.6 | 64.3 | 52.1 | 53.8 | 38.6 | 64.2 | 45.1 | 71.0 |
| Clusters | 36 | 36 | 36 | 36 | 36 | 35 | 35 | 30 | 30 |
| Registrants | 1,586,751 | 326,211 | 472,401 | 345,185 | 449,866 | 333,314 | 756,643 | 155,077 | 368,623 |
| Registrant-episode-elections | 6,505,033 | 1,422,470 | 1,842,629 | 1,320,973 | 1,918,961 | 1,395,521 | 3,200,212 | 592,077 | 1,317,223 |
| *Panel B: State chambers* | | | | | | | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.22*** | 2.00*** | 1.15*** | 1.64*** | 0.328 | 1.33*** | 1.02*** | 2.28*** | 0.963*** |
| | (0.243) | (0.307) | (0.319) | (0.294) | (0.324) | (0.495) | (0.273) | (0.800) | (0.279) |
| Turnout percentage | 58.0 | 61.9 | 63.6 | 51.8 | 53.3 | 38.7 | 63.6 | 44.2 | 69.6 |
| Clusters | 302 | 302 | 301 | 300 | 300 | 267 | 270 | 185 | 186 |
| Registrants | 4,653,157 | 948,498 | 1,391,972 | 1,088,385 | 1,336,279 | 1,096,844 | 2,186,641 | 576,787 | 1,117,965 |
| Registrant-episode-elections | 24,861,956 | 5,308,114 | 7,647,823 | 5,207,700 | 6,698,319 | 5,025,233 | 11,373,908 | 2,587,686 | 5,875,129 |

The table presents the coefficient estimates, $\hat{\alpha}_i^{\mathrm{USH}}$ and $\hat{\alpha}_i^{\mathrm{NC}}$, that are used in the simulation. Values are obtained by running Model (3) for the listed groups of registrants and types of chambers. The "common" specification uses the coefficient estimates from the "All" column; these match the values in the "Chamber" columns of Table 1. The "race-party" specification uses the estimates from the 2nd through 5th columns. The "age-education" specification uses those from the last four columns. Other details are the same as in Table 1.

Table A14: The change in registrants' turnout probabilities under 55-45 districts: robustness for the 2020 election

| **Specification** | **All registrants** | | | **Means by race and party** | | | | **Means by party** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | 10th percentile | 90th percentile | White-Dem. | White-Rep. | White-Unaffil. | Minority | Dem. | Rep. | Unaffil. |
| Race-party | 2.53 | 0.39 | 5.23 | 3.85 | 2.55 | 3.25 | 0.92 | 2.44 | 2.46 | 2.79 |
| Age-education | 2.67 | 0.56 | 4.73 | 2.42 | 2.46 | 2.61 | 3.15 | 2.78 | 2.47 | 2.73 |
| Common | 2.67 | 0.60 | 4.60 | 2.41 | 2.50 | 2.48 | 3.22 | 2.83 | 2.51 | 2.60 |

The table is analogous to the "2020 election" row of Table 6. "Race-party" is the specification used in the main text. "Age-education" lets $\hat{\alpha}_i^{\text{USH}}$ and $\hat{\alpha}_i^{\text{NC}}$ vary based on a registrant's age and the share college graduates in his baseline block-group. "Common" restricts $\hat{\alpha}_i^{\text{USH}}$ and $\hat{\alpha}_i^{\text{NC}}$ to be the same for all registrants. See Table A13 for the coefficient estimates used in these specifications.

Table A15: The change in aggregate turnout under 55-45 districts: robustness for the 2020 election

| **Specification** | **Actual turnout** | **Change in turnout** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **All** | **By race and party** | | | | **By party** | | |
| | | | White-Dem. | White-Rep. | White-Unaffil. | Minority | Dem. | Rep. | Unaffil. |
| Race-party | 3,467,293 | 143,393 | 48,651 | 43,445 | 37,033 | 14,264 | 59,711 | 44,168 | 39,514 |
| Age-education | 3,467,293 | 151,131 | 30,597 | 41,822 | 29,676 | 49,036 | 68,138 | 44,331 | 38,662 |
| Common | 3,467,293 | 151,255 | 30,380 | 42,569 | 28,176 | 50,130 | 69,239 | 45,111 | 36,905 |

The table is analogous to the "2020 election" row of Table 7. However, it provides results for alternative specifications.

Table A16: The change in aggregate votes under 55-45 districts: robustness for the 2020 election

| **Specification** | **Predicted votes** | | **Change in votes** | | **Net change for Dem.** |
|---|---|---|---|---|---|
| | Democrats | Republicans | Democrats | Republicans | |
| Race-party | 1,656,080 | 1,741,297 | 65,608 | 73,864 | -8,256 |
| Age-education | 1,656,080 | 1,741,297 | 80,658 | 66,456 | 14,202 |
| Common | 1,656,080 | 1,741,297 | 79,580 | 67,749 | 11,831 |

The table is analogous to the "2020 election" row of Table 8. However, it provides results for alternative specifications.

Table A17: The partisan composition of different groups' increases in votes: robustness for the 2020 election

| **Specification** | **By race and party** | | | | | | | | **By party** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | White-Dem. | | White-Rep. | | White-Una. | | Minority | | Dem. | | Rep. | | Unaffil. | |
| | Dem. | Rep. | Dem. | Rep. | Dem. | Rep. | Dem. | Rep. | Dem. | Rep. | Dem. | Rep. | Dem. | Rep. |
| Race-party | 64 | 34 | 14 | 85 | 44 | 50 | 83 | 15 | 69 | 29 | 14 | 85 | 45 | 49 |
| Age-education | 67 | 31 | 14 | 85 | 46 | 48 | 83 | 15 | 80 | 18 | 15 | 84 | 51 | 43 |
| Common | 64 | 34 | 14 | 85 | 44 | 50 | 83 | 15 | 79 | 19 | 15 | 84 | 49 | 45 |

The table is analogous to the "2020 election" row of Table 9. However, it provides results for alternative specifications.

# A1 Summary statistics

Summary statistics for the data are presented in Tables A18 and A19.

Table A18 lists the number of registrants in each of the baseline elections: 2010, 2014, 2016, and 2018. It reveals that the baseline elections included almost 9 million distinct registrants, with an average of 6.8 million registrants per election.

Table A18: The number of registrants in each baseline election

| **Baseline election** | **Registrants** |
|---|---|
| 2010 | 6,255,853 |
| 2014 | 6,664,171 |
| 2016 | 6,979,559 |
| 2018 | 7,152,496 |
| Average | 6,763,020 |
| Distinct | 8,963,975 |
| Total | 27,052,079 |

The table summarizes the number of registrants in the baseline elections. 2010 is the baseline for the decennial redistricting episodes. 2014 is the baseline for the first court-ordered revision for the U.S. House. 2016 is the baseline for the same revision for the NC Senate and NC House. 2018 is the baseline for episodes associated with the second court-ordered revision.

Panel A of Table A19 summarizes the data on registrants. Values are calculated using the registrant populations from the four baseline elections. On average, the registrants in these elections are 48 years old. 45% of them are male, with the others being female or unknown. 22% self-identify as black, 70% as white, and the remainder as other races. 41% register as Democrats, with 30% choosing to be Republicans and 29% staying unaffiliated. The registrants live in census block-groups where, on average, the median household income is $52,000 and where an average of 31% of adults are college graduates. The registrants' property parcels have a mean value of $83,500 per registered resident.

Panel B of Table A19 summarizes the data on legislative races. It reveals that only two-thirds of races during 2006 to 2020 were contested by both major parties. On average, the races had an absolute two-party vote-share margin of 49 percentage points. This gives them a mean closeness score of 0.51. In addition, the average spending in the races was $2.23 per district resident.

Finally, Panel C of the table shows that a fifth of the districts that were used in North Carolina during 2006 to 2020 were "majority-minority".

Table A19: Summary statistics

| | Mean | Std. dev. | N |
|---|---|---|---|
| *Panel A: Registrants* | | | |
| Demographics | | | |
| Age | 48.4 | 18.4 | 27,051,549 |
| Male | 0.452 | 0.498 | 27,052,079 |
| Black | 0.224 | 0.417 | 27,052,079 |
| White | 0.703 | 0.457 | 27,052,079 |
| Party registration | | | |
| Democrat | 0.407 | 0.491 | 27,052,079 |
| Republican | 0.304 | 0.460 | 27,052,079 |
| Unaffiliated | 0.289 | 0.453 | 27,052,079 |
| Census covariates | | | |
| Population density in census block | 1,171 | 2,824 | 27,052,079 |
| Median hhld. income in block-group (2010 $) | 51,948 | 25,801 | 26,676,434 |
| Share college graduates in block-group | 0.309 | 0.207 | 27,049,630 |
| Other covariates | | | |
| Parcel value per registrant (2010 $) | 83,461 | 375,139 | 26,067,323 |
| *Panel B: Legislative races* | | | |
| Contested by both parties | 0.673 | 0.469 | 1,464 |
| Closeness | 0.506 | 0.377 | 1,464 |
| Spending per person (2010 $) | 2.23 | 3.37 | 1,464 |
| *Panel C: Legislative districts* | | | |
| Majority-minority | 0.195 | 0.397 | 732 |

The table presents summary statistics for the data used in the paper. The sample in Panel A is registrants in the 2010, 2014, 2016, and 2018 elections. “Population density” is calculated as people per square km. “Share college graduates” is the fraction of adults age 25 and over who have graduated from college. “Parcel value per registrant” is calculated by dividing the value of a property parcel by the number of individuals registered at its address. The sample in Panel B is races for the U.S. House, NC Senate, and NC House between 2006 and 2020. “Contested by both parties” is an indicator for whether a race included candidates from both the Democratic and Republican parties. “Closeness” is 1 minus the absolute two-party vote-share margin. For uncontested races, it is equal to 0, since the two-party vote-share margin in these races is 1. “Majority-minority” is an indicator equal to 1 if more than 50% of a district’s registrants are non-white. This variable is calculated for all districts used between 2006 and 2020.

# A2 Additional details on the competitiveness measures

In this appendix, we provide additional details on the competitiveness measures. As discussed in the main text, we compute the measures in two steps. First, we predict district vote shares using the information available at the time of redistricting. Second, we define competitiveness as one minus a district's absolute predicted two-party vote-share margin. We create three different measures of competitiveness, which differ in the strategies that are used for predicting district vote shares.

## A2.1 Main measure

We calculate the main measure in four steps. First, in an episode's baseline election, we regress precinct-level vote shares on precinct-level means of voter characteristics. Let $\mathrm{v}_{hq}^{D}$ ($\mathrm{v}_{hq}^{R}$) be the Democratic (Republican) vote share in precinct $h$ on contest $q$. Let $\mathrm{X}_i$ be a vector of registrant $i$'s characteristics, and let $\mathrm{X}_h$ be the mean of these characteristics among registrants in $h$ who turned out to vote. We run regressions of the form:

$$\mathrm{v}_{hq}^{k} = \mathrm{X}_h'\eta^k + \eta_q^k + \eta_{hq}^k \quad \text{for } k \in \{D, R\}.$$

Here, $\eta_q^k$ is a contest fixed effect and $\eta_{hq}^k$ is an error term.[1]

Second, we use the regressions to generate individual-level predictions, $\mathrm{p}_{iq}^{k}$, of a registrant's probability of preferring party $k$ in contest $q$. These are obtained for all baseline registrants, not just those who turned out. We calculate the predictions as:

$$\mathrm{p}_{iq}^{k} = \mathrm{X}_i'\hat{\eta}^k + \hat{\eta}_q^k + \hat{\eta}_{hq}^k \quad \text{for } k \in \{D, R\}.$$

Here, a hat over a coefficient indicates that the value is a coefficient estimate. We bound the predictions between 0 and 1 if they fall outside the unit interval.

Third, we aggregate the predictions to the district level by computing weighted averages over all registrants whose baseline address is within a district's boundaries. We label the weighted averages for district $d$ as $\mathrm{v}_{dq}^{k}$. The weights depend on a registrant's turnout behavior in pre-redistricting elections. Specifically, we divide registrants into four groups based on turnout in the baseline and the election before the baseline. We calculate a weight for each group as the group's turnout rate in the five elections after the baseline. The weights ensure that the predicted vote shares, $\mathrm{v}_{dq}^{k}$, accumulate preferences in a way that reflects registrants' relative likelihood of voting in post-redistricting elections.[2]

1. The contests that are included in the regressions vary by baseline election, as elections differ in which contests are on the ballot. All baselines include contests for the U.S. House of Representatives. The 2010 and 2014 baselines add contests for U.S. Senate. 2016 adds U.S. President, U.S. Senate, NC Governor, the average of NC Attorney General and NC Secretary of State, and the average of other, more minor NC state offices. 2018 adds a NC Supreme Court seat and the average of three NC Court of Appeals seats. The regressions include numerous registrant characteristics: an indicator for self-identifying as black, an indicator for self-identifying as white, an indicator for being male, indicators for four age groups ($\leq 25$, $26-35$, $36-65$, and $\geq 66$), an indicator for leaning Democratic (being registered as a Democrat or being Unaffiliated but having voted in the Democratic primary), an indicator for leaning Republican (being registered as a Republican or being Unaffiliated but having voted in the Republican primary), an indicator for leaning Democratic and being at least 66 years old, parcel value per registrant (discretized into 20 groups), the natural log of block population density, the natural log of block-group median household income, and block-group share college graduates.

2. We choose five elections after the baseline for calculating turnout rates because legislative districts are meant to last for 10 years. We calculate the rates as the average of values for registrants from the 2008 and 2010 elections. These are the only elections for which we can observe both one prior election and five future elections. The weights

Finally, we calculate competitiveness as:

$$\mathrm{c}_{d,\mathrm{M}} = 1 - \left| \frac{1}{Q} \sum_q \frac{\mathrm{v}_{dq}^D - \mathrm{v}_{dq}^R}{\mathrm{v}_{dq}^D + \mathrm{v}_{dq}^R} \right|.$$

Here, the M subscript denotes that this is our main measure, the sum is over the contests in the baseline, and $Q$ is the number of such contests. In words, competitiveness for district $d$ is one minus the absolute value of the district's predicted two-party vote share margin, where the prediction is a mean of predictions that are based on each baseline contest.

### A2.2 Alternative measure

The second competitiveness measure is labeled, $\mathrm{c}_{d,\mathrm{A}}$. It makes one modification to the main measure. Namely, it does not use a regression to generate predictions for individual-level preference probabilities. Instead, it assigns each registrant a preference probability equal to the vote share in the registrant's precinct: i.e., $\mathrm{p}_{iq}^k = \mathrm{v}_{hq}^k$. In this way, the measure predicts district vote shares by calculating weighted averages of precinct vote shares. This captures how competitiveness could be measured if a researcher lacked data on registrant characteristics, other than turnout history.

### A2.3 Cook measure

The last competitiveness measure is built on the Partisan Voter Index (PVI) from the Cook Political Report. The PVI is commonly used both in academic literature (e.g., Moskowitz and Schneer (2019)) and by the media. It predicts a district's vote shares based on how the district's residents voted in past presidential races. The Cook Political Report provides the PVI for the decennial U.S. House districts but not for the revised U.S. House districts or for the state legislative districts. As such, we calculate the PVI ourselves.

The PVI is the difference in the two-party presidential vote share between the district and the entire country, averaged over two recent elections. It is:

$$\mathrm{PVI}_d = \frac{1}{2} \sum_{p \in \{p_1, p_2\}} \left( \frac{\mathrm{v}_{dp}^D}{\mathrm{v}_{dp}^D + \mathrm{v}_{dp}^R} - \frac{\mathrm{v}_p^D}{\mathrm{v}_p^D + \mathrm{v}_p^R} \right).$$

Here, $\mathrm{v}_{dp}^k$ is the district-$d$ vote share in presidential race $p$ for party $k$, and $\mathrm{v}_p^k$ is the nationwide vote share for this party. In our implementation, we choose $p_1$ and $p_2$ to be the races from the first and second presidential elections prior to a redistricting episode. We calculate district vote shares by aggregating precinct votes to the district level.[3] Due to data limitations, we cannot do the aggregation for the 2004 presidential race. Thus, in situations where 2004 is required, we use only the vote from 2008, the first presidential election prior to the decennial redistricting.

The PVI is a prediction of partisan lean, not the two-party vote-share margin. In particular, it predicts the difference between a district's two-party Democratic vote share and the national two-party Democratic vote share (which is a value close to 0.5). This difference is about half as large as the difference between the district's two-party Democratic and Republican vote shares.

are: 0.79 for registrants who voted in both the baseline and the election before the baseline, 0.57 for registrants who voted only in the baseline, 0.37 for registrants who voted only in the election before the baseline, and 0.13 for registrants who didn't vote in either election.

3. For the aggregation, we use a regression to predict individual-level preference probabilities, as in the procedure for our main measure. We then compute an unweighted average of the predictions over all registrants in a district's boundaries who voted in the election. This way, we are merely aggregating the observed vote, not adjusting it based on turnout probabilities, as we do with the other measures.

Thus, in order to be consistent with the other competitiveness measures, we multiply the PVI by two when we calculate the Cook measure. It is:

$$c_{d,\text{PVI}} = 1 - |2 \cdot \text{PVI}_d|.$$

Under this construction, a district that votes 50 percentage points more partisan than the country as a whole would have $c_{d,\text{PVI}} = 0$; meanwhile, one that votes the same as the entire country would have $c_{d,\text{PVI}} = 1$. Provided that the nationwide presidential vote is about even, on average, over the two elections used in calculating the PVI, then these districts would be close to 100-0 and 50-50 in terms of predicted two-party vote shares. Thus, the scaling for the Cook measure is similar to that for our other measures.

## A2.4 Primary v. secondary versions

We compute versions of district competitiveness for each baseline election. For a given district, the "primary" version is calculated during the baseline for the redistricting episode that created the district. By contrast, "secondary" versions are calculated during the baselines for other episodes. The primary version reflects our knowledge of a district's competitiveness when the district was being drawn. The secondary versions are important for our empirical strategy. Namely, in the causal analysis, we want to observe competitiveness at the time of a given episode's baseline both for the districts created by the episode and for districts that are used in earlier or later elections. Values for these latter districts are captured by the secondary versions of the measures.

## A2.5 Summary statistics and prediction quality

Summary statistics for the primary versions of the three competitiveness measures are presented in Table A20; a correlation matrix is provided in Table A21. The tables show that the measures have similar means and standard deviations and are highly correlated. However, the alternative measure has a slightly smaller mean and a slightly larger standard deviation. Also, the Cook measure is less correlated with the main and alternative measures (0.89 and 0.91) than those measures are with each other (0.97).

Table A20: Summary statistics for the competitiveness measures

| | **Mean** | **Std. dev.** | **N** |
|---|---|---|---|
| Main measure, $c_{d,\text{M}}$ | 0.766 | 0.139 | 549 |
| Alternative measure, $c_{d,\text{A}}$ | 0.721 | 0.159 | 549 |
| Cook measure, $c_{d,\text{PVI}}$ | 0.762 | 0.127 | 549 |

The table presents summary statistics for the three measures of district competitiveness. The sample is the 549 districts that were used in North Carolina during the 2012-2020 elections. The table uses the primary versions of the measures (i.e., calculated during the baseline elections for the districts' own redistricting episodes).

Table A21: Correlations among measures of district competitiveness

| | Main measure | Alt. measure | Cook measure |
|---|---|---|---|
| Main measure, $c_{d,\text{M}}$ | 1 | - | - |
| Alternative measure, $c_{d,\text{A}}$ | $0.971^{***}$ | 1 | - |
| Cook measure, $c_{d,\text{PVI}}$ | $0.890^{***}$ | $0.909^{***}$ | 1 |

The table presents a correlation matrix for the three measures of district competitiveness. See Table A20 for details on the sample and measures.

Table A22 presents results regarding the predictive power of the measures. Similar to Figure 2, it lists coefficient estimates, standard errors, and R-squared for regressions of outcomes in legislative races on measures of the competitiveness of the races' districts. The race outcomes are race closeness, an indicator for whether the race was contested by both parties, and race spending. The columns in the table are for the three competitiveness measures. The panels are for different samples, with Panel A being all legislative races and the other panels being races in only the specified chambers. In terms of findings, the table reveals that all measures have predictive power, but our main measure has the most and the Cook measure has the least.

Table A22: Predicting race outcomes using district competitiveness measures: results by legislative chamber

| | **Main measure** | | **Alt. measure** | | **Cook measure** | |
|---|---|---|---|---|---|---|
| | Coef. (s.e.) | R-sq. | Coef. (s.e.) | R-sq. | Coef. (s.e.) | R-sq. |
| *Panel A: All chambers* (N=915) | | | | | | |
| Closeness | 1.58 (0.07) | 0.35 | 1.31 (0.07) | 0.30 | 1.56 (0.08) | 0.28 |
| Contested by both parties | 1.17 (0.10) | 0.13 | 0.89 (0.09) | 0.09 | 1.03 (0.12) | 0.08 |
| Ln. spending per person | 5.05 (0.25) | 0.32 | 4.15 (0.23) | 0.27 | 5.06 (0.28) | 0.26 |
| *Panel B: U.S. House* (N=65) | | | | | | |
| Closeness | 1.23 (0.19) | 0.41 | 1.34 (0.18) | 0.46 | 1.56 (0.31) | 0.29 |
| Contested by both parties | 0.27 (0.24) | 0.02 | 0.37 (0.24) | 0.04 | 0.02 (0.36) | 0.00 |
| Ln. spending per person | 3.65 (0.95) | 0.19 | 3.78 (0.96) | 0.20 | 4.14 (1.48) | 0.11 |
| *Panel C: NC Senate* (N=250) | | | | | | |
| Closeness | 1.57 (0.14) | 0.33 | 1.29 (0.13) | 0.28 | 1.53 (0.16) | 0.26 |
| Contested by both parties | 1.03 (0.20) | 0.09 | 0.75 (0.18) | 0.06 | 0.87 (0.23) | 0.06 |
| Ln. spending per person | 5.61 (0.49) | 0.35 | 4.48 (0.46) | 0.28 | 5.61 (0.56) | 0.29 |
| *Panel D: NC House* (N=600) | | | | | | |
| Closeness | 1.56 (0.09) | 0.34 | 1.28 (0.08) | 0.30 | 1.52 (0.10) | 0.27 |
| Contested by both parties | 1.21 (0.13) | 0.13 | 0.90 (0.11) | 0.09 | 1.04 (0.14) | 0.08 |
| Ln. spending per person | 4.96 (0.30) | 0.31 | 4.05 (0.27) | 0.27 | 4.91 (0.34) | 0.26 |

The table presents results from regressions of outcomes in legislative races on the competitiveness of the races' districts. The outcomes are listed in the rows of the table. The column titled "Main measure" is for regressions that use the main competitiveness measure, $c_{d,\mathrm{M}}$. The results under "Alt. measure" and "Cook measure" instead use the alternative and Cook measures, $c_{d,\mathrm{A}}$ and $c_{d,\mathrm{PVI}}$. All competitiveness measures are calculated during the baseline election for a district's own redistricting episode. The sample in Panel A is all legislative races that occurred during the 2012-2020 elections. The samples in the other panels are restricted to races in the specified chambers. Sample sizes are in parenthesis in the panel headings. "Contested by both parties" is an indicator for whether a race features both a Democratic and Republican candidate. See Figure 2 for definitions of the other outcomes.

# A3 Additional details on the matching procedure

In this appendix, we present additional details on the matching procedure. We first explain the construction of regions. We then discuss the covariates that are used in matching. Finally, we describe the estimation sample.

## A3.1 Regions in the main analysis

In the main analysis, regions are meant to identify registrants who had similar district experiences in (i) pre-redistricting elections for all chambers and (ii) post-redistricting elections for chambers other than the chamber of interest. As such, we construct regions as intersections of districts that were used before and after redistricting, but excluding those used after redistricting for the chamber of interest. In matching, we focus on the region that a registrant lives in during the baseline election. Importantly, the registrant may not experience all the districts that are used in constructing this region, as she may live in different regions in non-baseline elections. However, she does experience these districts if she does not move.

Table A23: The districts used in constructing regions

| **Episode** | **Pre-2011** | | | **Decennial** | | | **1st revision** | | | **2nd revision** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | USH | NCS | NCH | USH | NCS | NCH | USH | NCS | NCH | USH | NCS | NCH |
| Decennial redistricting | | | | | | | | | | | | |
| U.S. House | Y | Y | Y | | Y | Y | | | | | | |
| NC Senate | Y | Y | Y | Y | | Y | | | | | | |
| NC House | Y | Y | Y | Y | Y | | | | | | | |
| 1st court-ordered revision | | | | | | | | | | | | |
| U.S. House | Y | Y | Y | Y | Y | Y | | Y | Y | | | |
| NC Senate | Y | Y | Y | Y | Y | Y | Y | | Y | | | |
| NC House | Y | Y | Y | Y | Y | Y | Y | Y | | | | |
| 2nd court-ordered revision | | | | | | | | | | | | |
| U.S. House | | | | Y | Y | Y | Y | Y | Y | | Y | Y |
| NC Senate | | | | Y | Y | Y | Y | Y | Y | Y | | Y |
| NC House | | | | Y | Y | Y | Y | Y | Y | Y | Y | |

The table shows the districts that are used in constructing the regions in the paper's main analysis. A "Y" indicates that the given districts are used for the regions for the specified redistricting episode.

Table A23 lists the districts that are used in constructing the regions for each episode. Regions for the decennial redistricting are the intersection of the pre-2011 districts for all chambers and the decennial districts for chambers other than the chamber of interest. Regions for the first court-ordered revision are the intersection of the pre-2011 and decennial districts for all chambers and the first-revision districts for chambers other than the chamber of interest. Finally, regions for the second court-ordered revision are the intersection of the decennial and first-revision districts for all chambers and the second-revision districts for chambers other than the chamber of interest.[4]

Table A24 displays the number of regions by episode. This varies from a low of 659 for the decennial episode for the NC House to a high of 1,452 for the first revision for the U.S. House. In comparison, North Carolina has almost 2,200 census tracts. Thus, regions can be understood as slightly larger than a census tract.

4. In this construction, decennial regions share the same districts for four elections before redistricting (for all chambers) and for two after (for chambers other than the chamber of interest). Values for the first-revision for the U.S. House (state chambers) are six (seven) before and two (one) after. For the second-revision, values are four before and one after.

Table A24: The number of regions by episode

| Episode | Regions |
|---|---|
| Decennial redistricting | |
| U.S. House | 900 |
| NC Senate | 833 |
| NC House | 659 |
| 1st court-ordered revision | |
| U.S. House | 1,452 |
| NC Senate | 1,424 |
| NC House | 1,298 |
| 2nd court-ordered revision | |
| U.S. House | 1,064 |
| NC Senate | 1,039 |
| NC House | 958 |
| All episodes | 9,627 |

The table displays the number of regions by redistricting episode. See Appendix A3.1 for details on the construction of regions.

## A3.2 District variables used in robustness checks

In robustness checks, we attempt to account for the fact that registrants may move to different regions in non-baseline elections. We do this by matching on additional district variables. In one specification ("Own district in election $-2$"), we match both on a registrant's baseline region and on her chamber-of-interest district in election $\tau = -2$. This election is the first before the baseline. In another specification ("Own district in elections $-2$ and $-3$"), we match on a registrant's baseline region and on her chamber-of-interest districts in both $\tau = -2$ and $\tau = -3$. In a third specification ("Own districts"), we drop regions and instead match on a set of the registrant's districts in pre- and post-redistricting elections. This set is the registrant's districts in all chambers in the three elections before redistricting and the registrant's districts in chambers other than the chamber of interest in the first election after redistricting.

## A3.3 Regions in the additivity analysis

After conducting the main analysis, we explore how the effects of district competitiveness aggregate across chambers. This requires us to compare registrants who differ in assigned districts for multiple chambers. To do so, we alter the definition of regions by not incorporating districts used in post-redistricting elections. Specifically, in the decennial redistricting, regions are the intersection of the pre-2011 districts. In the first revision for the U.S. House, regions are the intersection of the pre-2011 and decennial districts. In the first revision for the NC House and NC Senate, regions are the intersection of the pre-2011 and decennial districts for all chambers and the first revision districts for the U.S. House.[5] Finally, in the second revision, regions are the intersection of the decennial and first-revision districts for all chambers.

## A3.4 The covariates used in matching

Our main specification uses the following covariates in matching: gender, three race/ethnicity groups, five age groups, three groups for the share college graduates in the registrant's baseline block-group, three groups for the registrant's party registration in the baseline election, the registrant's history of turnout in the three elections prior to redistricting, and the election in which

5. We include the first revision districts for the U.S. House because this episode occurred one election before the first revision for the state chambers.

the registrant first registered in North Carolina.

Gender is defined as male or other. Race/ethnicity is defined as white, black, or other. Age is measured in the baseline election, and the five age groups are $\leq 25$, $26-35$, $36-50$, $51-65$, and $\geq 66$. Registrants with missing age are joined with the $\geq 66$ group. The three groups for share college graduates are $\leq 0.2$, $0.2-0.4$, and $> 0.4$. Before discretizing this variable, we impute missing values using tract- and county-level medians. The three groups for party registration are Democrat, Republican, and Unaffiliated, with third parties counted as Unaffiliated. The registrant's history of turnout in the three elections prior to redistricting is captured by a set of three variables that indicate whether the registrant turned out in each election.

In robustness checks, we add variables related to: (i) the value of the registrant's baseline property parcel, (ii) the population density of the registrant's baseline census block, and (iii) the median household income in the registrant's baseline block-group. For parcel value, we use an indicator for whether the per-registrant value of the property parcel is at least \$100,000 in 2010 dollars. For population density, we use an indicator for whether the block has more than 1,500 people per square kilometer. For household income, we use an indicator for whether the block-group's median household income is greater than \$50,000 (again in 2010 dollars). Finally, as discussed in Appendix A3.2, we sometimes match on additional district variables.

## A3.5 Characterizing the main estimation sample

We now characterize the main estimation sample. This is the sample associated with our main specification of regions and covariates. It is for the analysis where we study registrants who differ in assigned post-redistricting districts for only a single chamber.

We create the sample in three steps. First, we limit attention to registrants in a baseline election who do not die before the 2020 election. Second, we run the matching procedure. Third, we drop registrants who are in match-groups with no variation in the assigned district.

The size of the estimation sample is shown in Table A25. The sample includes almost 9 million registrant-episode combinations. These are in almost 1,300 regions and over 500,000 match-groups. On average, there are 7,000 registrants per region and 17 registrants per match-group.

Table A26 presents summary statistics on covariates for the estimation sample. It reveals that means and standard deviations are similar to those for all North Carolina registrants (Table A19).

Figure A4 provides a graphical illustration of the regions that are used in generating the estimation sample. It plots regions for the decennial episode for the NC House for two example counties. The first (Wake) is the largest county in the state and is home to Raleigh, the state capital. The second (Buncombe) is a moderately sized county and is home to Asheville.

Finally, Figure A5 summarizes the locations in North Carolina that contribute to the estimation sample. The area colored in blue reveals all census block-groups that contain estimation-sample registrants during baseline elections. The figure shows that registrants come from across the state.

Table A25: The size of the estimation sample

| Episode | Registrants | Regions | Registr. per region | | Match-groups | Registr. per m.-g. | |
|---|---|---|---|---|---|---|---|
| | | | Mean | Std. dev. | | Mean | Std. dev. |
| Decennial redistricting | | | | | | | |
| U.S. House | 1,047,606 | 190 | 5,514 | 6,593 | 56,092 | 18.7 | 33.3 |
| NC Senate | 1,193,592 | 255 | 4,681 | 7,132 | 63,741 | 18.7 | 36.4 |
| NC House | 2,894,412 | 336 | 8,614 | 11,479 | 125,108 | 23.1 | 48.5 |
| 1st court-ordered revision | | | | | | | |
| U.S. House | 307,022 | 54 | 5,686 | 6,238 | 18,962 | 16.2 | 36.1 |
| NC Senate | 329,706 | 68 | 4,849 | 5,459 | 24,770 | 13.3 | 32.2 |
| NC House | 1,111,225 | 183 | 6,072 | 7,457 | 81,701 | 13.6 | 33.0 |
| 2nd court-ordered revision | | | | | | | |
| U.S. House | 345,937 | 32 | 10,811 | 13,391 | 24,118 | 14.3 | 37.4 |
| NC Senate | 387,344 | 51 | 7,595 | 7,325 | 31,011 | 12.5 | 27.6 |
| NC House | 1,152,730 | 116 | 9,937 | 12,705 | 82,442 | 14.0 | 32.8 |
| All episodes | 8,769,574 | 1,285 | 6,825 | 9,295 | 507,945 | 17.3 | 38.0 |

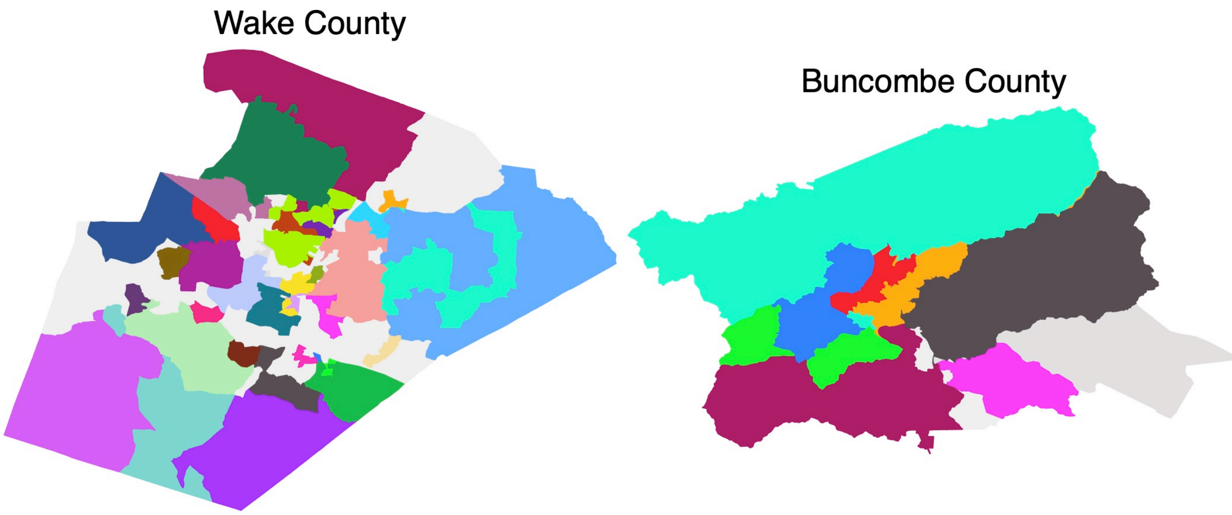
The table describes the size of the estimation sample for our main specification. The estimation sample draws from individuals registered in a baseline election. It excludes registrants who die before the 2020 election and registrants who are in match-groups with no variation in the assigned post-redistricting district. For the decennial redistricting episodes, the baseline election is 2010. For episodes from the 1st court-ordered revision, the baseline election is 2014 for the U.S. House and 2016 for the NC Senate and NC House. For episodes from the 2nd court-ordered revision, the baseline election is 2018. In the row labeled “All episodes”, the value under “Registrants” is the number of combinations of registrants and redistricting episodes. The estimation sample includes 5,203,371 distinct registrants. “Registr. per region” and “Registr. per m.-g.” are, respectively, the number of registrants per region and match-group. See Appendices A3.1, A3.4, and A3.5 for more details.

Table A26: Summary statistics for the estimation sample

| | Mean | Std. dev. | N |
|---|---|---|---|
| Demographics | | | |
| Age | 46.4 | 17.1 | 8,769,526 |
| Male | 0.441 | 0.497 | 8,769,574 |
| Black | 0.225 | 0.418 | 8,769,574 |
| White | 0.723 | 0.447 | 8,769,574 |
| Party registration | | | |
| Democrat | 0.421 | 0.494 | 8,769,574 |
| Republican | 0.310 | 0.462 | 8,769,574 |
| Unaffiliated | 0.269 | 0.443 | 8,769,574 |
| Census covariates | | | |
| Population density in census block | 1,237 | 2,293 | 8,769,574 |
| Median hhld. income in block-group (2010 $) | 55,390 | 28,480 | 8,705,170 |
| Share college graduates in block-group | 0.334 | 0.221 | 8,769,139 |
| Other covariates | | | |
| Parcel value per registrant (2010 $) | 87,671 | 407,829 | 8,493,779 |

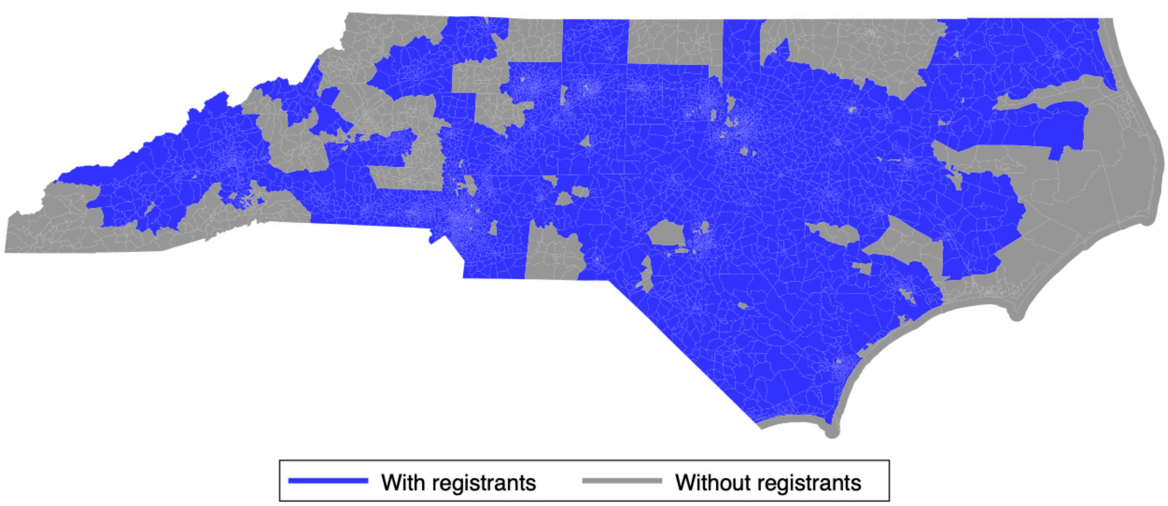
The table presents summary statistics for the registrants included in the main estimation sample. See Table A25 for details on the size and construction of the estimation sample. See Table A19 and Appendix A3.4 for definitions of covariates.

Figure A4: Regions for the decennial episode for the NC House




The figure provides examples of the regions that are used in obtaining the main estimation sample. Areas with different colors represent different regions. All regions that lack estimation-sample registrants are shaded light gray. The regions are for the decennial redistricting episode for the NC House. The left panel depicts all regions that overlap with Wake County. The right panel depicts all regions that overlap with Buncombe County. See Appendix A3.1 for details on the construction of regions and Appendix A3.5 for details on the estimation sample.

Figure A5: Census block-groups with registrants in the estimation sample




The figure reveals the geographic distribution of the registrants in the main estimation sample. The area colored in blue is the census block-groups in which estimation-sample registrants lived during the baseline elections. See Appendix A3.5 for more details.

# A4 The first stage in the IV model

In this appendix, we discuss the first stage in the IV model. The first stage is the relationship between the treatment variable and the instruments. In Models (2) and (3), it is represented by the equation:

$$\mathrm{C}_{i\tau} = \beta_\tau \cdot \mathrm{C}_{a_i\tau} + \beta_{g_i\tau} + \beta_{i\tau}$$

This specification contains multiple instruments, given that it allows $\beta$ to vary by relative election. The instruments are the interaction of $\mathrm{C}_{a_i\tau}$ and indicators for $\tau$.

We next explain why we choose these variables as our instruments and we consider some alternatives. We then show that our setting has a strong first stage. Finally, we present a few mathematical results on the structure of the first-stage coefficients.

## A4.1 Choosing instruments

To ground our choice of instruments, we start by illustrating the first stage graphically. For each redistricting episode, we show how the episode creates variation in the competitiveness that registrants experience.

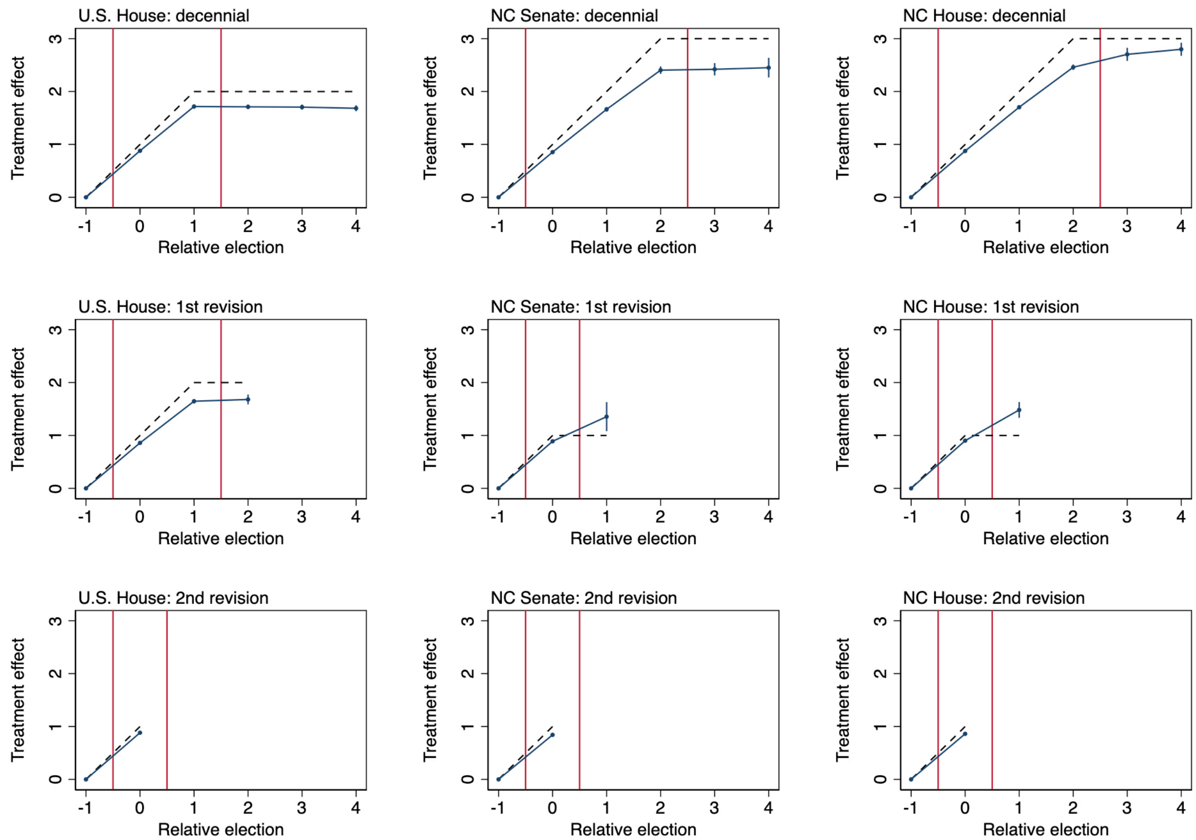
Figure A6 presents the results. In the figure, each plot depicts a different redistricting episode. The vertical red lines demarcate the relative elections during which the districts from the specified episode are in use. The solid blue lines are our quantities of interest. For each episode, they display estimates for the coefficients on assigned competitiveness, $\mathrm{c}_{a_i}$, in $\tau$-specific regressions of $\mathrm{C}_{i\tau}$ on $\mathrm{c}_{a_i}$ and match-group fixed effects. These regressions are episode-specific versions of Equation (1). The coefficients capture the effects on $\mathrm{C}_{i\tau}$ of being assigned to a 50-50 district versus a 100-0 district in the episode.

As a comparison, each of the plots also includes a dashed black line. These lines portray analogous effects as above, but on $\mathrm{C}_{a_i\tau}$ instead of $\mathrm{C}_{i\tau}$. The effects on $\mathrm{C}_{a_i\tau}$ show what would happen if (i) no registrants move out of their assigned districts before the next redistricting episode and (ii) $\mathrm{c}_{a_i}$ has no predictive power—within match-groups—for the competitiveness that registrants experience after subsequent episodes. The effects have a simple formula, owing to the mechanical relationship between $\mathrm{C}_{a_i\tau}$ and $\mathrm{c}_{a_i}$. Namely, they equal $\tau + 1$ in elections in which an episode's districts are still in use (i.e., for $\tau \leq \tau^l$) and $\tau^l + 1$ in all later elections (i.e., for $\tau > \tau^l$). Finally, the plots enable calculating episode-specific versions of the first-stage coefficients, $\beta_\tau$. These are equal to the ratio of a given effect on $\mathrm{C}_{i\tau}$ to the same effect on $\mathrm{C}_{a_i\tau}$.[6]

The plots in Figure A6 illuminate the mechanics behind the redistricting episodes. Notably, they reveal how the episodes lead to differences in $\mathrm{C}_{i\tau}$ within match-groups and how these differences change over time. For instance, for the decennial episode for the U.S. House, the effect of $\mathrm{c}_{a_i}$ on $\mathrm{C}_{i\tau}$ grows for the two elections in which the districts are in use and then is stable thereafter. In other words, being assigned to a more competitive district in this episode causes a registrant to experience a higher degree of competitiveness than others in his match-group for two elections and then has no effect in the remaining elections. The mechanics of the decennial episode for the NC Senate are similar; however, for this episode, the effect on $\mathrm{C}_{i\tau}$ grows for three elections (mirroring that the districts for this episode last for three elections). By contrast, the story for the decennial episode for the NC House is somewhat different. For this episode, the effect continues to grow—by a small amount—after the districts are revised. That is, individuals assigned to more competitive districts experience higher competitiveness for three elections and then also experience slightly higher competitiveness in elections after the next episode. This suggests that

6. The simple formula arises because $\mathrm{C}_{a_i\tau}$ is collinear with $\mathrm{c}_{a_i}$ for a given combination of episode and $\tau$.

Figure A6: The effect of assigned competitiveness, $c_{a_i}$, on $C_{i\tau}$, by episode




The figure presents a graphical illustration of the first stage. In the figure, each plot is a different redistricting episode. In the plots, the left red line designates the specified episode, and the right red line designates the next episode. Thus, the districts for the specified episode are in use during the elections between the two red lines. The solid blue lines display coefficient estimates for the coefficients on $c_{a_i}$ in episode-by-$\tau$-specific regressions of $C_{i\tau}$ on $c_{a_i}$ and match-group fixed effects. The dashed black lines reveal corresponding coefficient estimates from regressions that use $C_{a_i\tau}$ as the outcome variable. The vertical bars represent 95% confidence intervals. Episode-specific first-stage coefficients are equal to the ratios of the solid blue lines to the dashed black lines. Standard errors are clustered by baseline district in the chamber of interest.

policymakers in the subsequent episodes maintained some features of the decennial districts when drawing the new ones. The mechanics of the other episodes in Figure A6 can be interpreted in an analogous manner.

Another takeaway from Figure A6 is that the effects on $C_{a_i\tau}$ follow a similar pattern as those on $C_{i\tau}$. Put differently, the episode-specific first-stage coefficients are all close in magnitude. This is useful because it suggests that we can obtain a strong first stage without needing to include a large number of instruments. In particular, we aren't forced to allow the first-stage coefficients to vary by relative election and episode. Instead, we may be able to predict $C_{i\tau}$ almost as well if we specify them to vary only by relative election. Reducing the number of instruments eases the computational burden that we face in fitting IV models with tens of millions of observations.

Consistent with the previous discussion, in our main specification we choose the first-stage coefficients to vary only by relative election. That is, we select the instruments to be the interaction of $C_{a_i\tau}$ and indicators for $\tau$. Nonetheless, we also demonstrate that IV estimates are robust to using three alternative sets of instruments. These are: (i) just $C_{a_i\tau}$, (ii) the interaction of $C_{a_i\tau}$ and indicators for treatment-group-by-$\tau$, and (iii) the interaction of $C_{a_i\tau}$ and indicators for episode-by-$\tau$.[7] In the next subsection, we provide evidence that the first stages associated with these instruments are all strong.

## A4.2 The strength of the first stage

First-stage results for the specification used in Models (2) and (3) are presented in Table A27. The instruments in these models are interactions of $C_{a_i\tau}$ with indicators for $\tau$. Results are obtained by running $\tau$-specific regressions of the treatment, $C_{i\tau}$, on $C_{a_i\tau}$ and match-group fixed effects.

Table A27: The first stage in Models (2) and (3)

| | **Election relative to redistricting, $\tau$** | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.872*** | 0.884*** | 0.818*** | 0.871*** | 0.893*** |
| | (0.004) | (0.012) | (0.007) | (0.017) | (0.019) |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results from the first-stage specification used in Models (2) and (3). Specifically, it shows coefficient estimates and standard errors for $\beta_\tau$ from $\tau$-specific regressions of $C_{i\tau}$ on $C_{a_i\tau}$ and match-group fixed effects. The columns display the results for the specified election. "Registrant-episodes" is the number of observations. Standard errors are clustered by baseline district in the chamber of interest. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results in Table A27 indicate that the first stage is strong. Coefficient estimates for $\beta_\tau$ vary between 0.82 and 0.89, while the maximum standard error is merely 0.019.

Table A28: The first stage for a specification with a single instrument

| | **All** |
|---|---|
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.864*** |
| | (0.011) |
| Clusters | 338 |
| Registrants | 5,203,371 |
| Registrant-episode-elections | 31,366,989 |

The table presents results from a first-stage specification with a single instrument, $C_{a_i\tau}$. Specifically, it shows the coefficient estimate and standard error for the coefficient on $C_{a_i\tau}$ in a regression of $C_{i\tau}$ on $C_{a_i\tau}$ and match-group-by-$\tau$ fixed effects. "Registrant-episode-elections" is the number of observations. All other details are the same as in Table A27.

Tables A28 and A29 provide results for two of the alternative sets of instruments: only $C_{a_i\tau}$ (Table A28) and interactions of $C_{a_i\tau}$ with indicators for treatment-group-by-$\tau$ (Table A29). The results in these tables again reveal that the first stages are strong. In addition, the coefficient estimates are similar to those in Table A27.

The last set of instruments are the interactions of $C_{a_i\tau}$ with indicators for episode-by-$\tau$. We don't provide a table of results for this specification because it would have a large number of values. However, as discussed in Appendix A4.1, information on the specification's first stage can be gleaned from Figure A6. The figure suggests that the first stage is again strong, as the coefficient estimates in the plots are all much bigger than their confidence intervals. Further, the figure indicates that the first-stage coefficients are usually slightly less than 1, just as with the other specifications.

7. The first instrument can be used to estimate $\alpha$ but is insufficient for estimating $\alpha_\tau$. Also, as in Figure A1, "treatment groups" are sets of episodes whose districts last for the same length of time.

Table A29: The first stage by treatment group

| | Election relative to redistricting, $\tau$ | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| *Group A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.869*** | 0.846*** | 0.814*** | 0.873*** | 0.898*** |
| | (0.006) | (0.007) | (0.007) | (0.018) | (0.020) |
| Clusters | 138 | 138 | 138 | 138 | 138 |
| Registrants | 3,415,768 | 3,415,768 | 3,415,768 | 3,415,768 | 3,415,768 |
| Registrant-episode-elections | 4,088,004 | 4,088,004 | 4,088,004 | 4,088,004 | 4,088,004 |
| *Group B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.875*** | 0.851*** | 0.852*** | 0.853*** | 0.842*** |
| | (0.008) | (0.009) | (0.010) | (0.012) | (0.014) |
| Clusters | 25 | 25 | 25 | 13 | 13 |
| Registrants | 1,303,348 | 1,303,348 | 1,303,348 | 1,047,606 | 1,047,606 |
| Registrant-episode-elections | 1,354,628 | 1,354,628 | 1,354,628 | 1,047,606 | 1,047,606 |
| *Group C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.874*** | 1.44*** | - | - | - |
| | (0.006) | (0.071) | - | - | - |
| Clusters | 175 | 92 | - | - | - |
| Registrants | 2,791,776 | 1,352,931 | - | - | - |
| Registrant-episodes | 3,326,942 | 1,440,931 | - | - | - |

The table presents results from a first-stage specification in which the instruments vary by treatment group and relative election. Specifically, it shows coefficient estimates and standard errors for the coefficients on $C_{a_i\tau}$ in treatment-group-by-$\tau$-specific regressions of $C_{i\tau}$ on $C_{a_i\tau}$ and match-group fixed effects. The rows present the results for the specified treatment group. The columns display the results for the specified election. Group A is the redistricting episodes in which districts last for three elections. Group B (C) is the episodes in which districts last for two elections (one election). Other details are as in Table A27.

In sum, our setting has a strong first stage. This holds for multiple sets of instruments.

## A4.3 Mathematical results for the first-stage coefficients

We now present a few mathematical results regarding the first-stage coefficients. These yield a deeper understanding of how the instruments generate variation in $C_{i\tau}$. Importantly, the results are derived only for versions of the coefficients that vary by episode. Coefficients that instead average over multiple episodes may exhibit slightly different features—due to differences across the episodes in how long the districts last and in which relative elections have non-missing data.

To understand the results, consider an episode-specific first-stage coefficient. This is $\beta_\tau$ in a version of the first stage, $C_{i\tau} = \beta_\tau \cdot C_{a_i\tau} + \beta_{g_i\tau} + \beta_{i\tau}$, that uses data from only a single episode. For this coefficient, we have four notable results. First, in one special case, $\beta_\tau$ will equal 1. This is if no registrants move out of their assigned districts and if the districts are still in use in $\tau$. Second, if some registrants move out between the baseline election and $\tau$, then $\beta_\tau$ will be less than 1. Third, $\beta_\tau$ will decline over time, as more people leave their assigned districts. Fourth, this decay is likely to stop once there is a subsequent redistricting episode—that is, for $\tau > \tau^l$, $\beta_\tau$ will not necessarily be smaller than $\beta_{\tau^l}$. This is because $C_{a_i\tau}$ remains constant in elections after $\tau^l$. In fact, in these elections, $\beta_\tau$ may even be larger than $\beta_{\tau^l}$, which will occur if the districts drawn in the subsequent episodes resemble the earlier ones.

We derive the above-described results in three steps. First, we rewrite $\beta_\tau$ in a convenient way. Second, we make two simplifying assumptions. Third, we show that the assumptions, together with the setting, lead to the results.

### A4.3.1 Rewriting $\beta_\tau$

To rewrite $\beta_\tau$, we use the following manipulation. First, we eliminate the fixed effects in the first-stage equation by de-meaning the variables. We get:

$$\mathrm{C}_{i\tau} - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau] = \beta_\tau \cdot (\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]) + \beta_{i\tau}.$$

Second, we use the formula for an OLS coefficient to obtain a formula for $\beta_\tau$. We get:

$$\beta_\tau = \frac{\mathrm{E}[(\mathrm{C}_{i\tau} - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau]) \cdot (\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau])|\tau]}{\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau])^2|\tau]}.$$

This expression reveals that $\beta_\tau$ is a simple ratio. The denominator is the within-match-group variance of $\mathrm{C}_{a_i\tau}$ in election $\tau$. The numerator is the within-match-group covariance of $\mathrm{C}_{a_i\tau}$ and $\mathrm{C}_{i\tau}$ in this election.

Finally, we use the law of iterated expectations to adjust the formula for $\beta_\tau$. We get:

$$\begin{aligned}\beta_\tau &= \frac{\mathrm{E}[\mathrm{E}\{(\mathrm{C}_{i\tau} - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau]) \cdot (\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau])|g_i,\tau,\mathrm{c}_{a_i}\}|\tau]}{\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau])^2|\tau]} \\ &= \frac{\mathrm{E}[(\mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau,\mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau]) \cdot (\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau])|\tau]}{\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau])^2|\tau]}. \qquad (5)\end{aligned}$$

This last expression is convenient because it gives us a strategy for assessing the magnitude of $\beta_\tau$. Namely, we can do so by manipulating $\mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau,\mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau]$ and $\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]$.

### A4.3.2 Simplifying assumptions

In order to derive the claims from Appendix A4.3, we make two simplifying assumptions.

First, we assume that, within each match-group, a registrant's assigned competitiveness does not affect the registrant's probability of being registered in her assigned district, $a_i$. Formally, let $s_{i\tau}$ be an indicator for whether registrant $i$ is still registered in $a_i$ in election $\tau$. We assume that this variable is mean-independent of $\mathrm{c}_{a_i}$, conditional on match-groups and $\tau$. I.e.,

$$\Pr[s_{i\tau} = 1|g_i,\tau,\mathrm{c}_{a_i}] = \Pr[s_{i\tau} = 1|g_i,\tau]. \qquad (6)$$

Second, we assume that when (i) a registrant leaves her assigned district and (ii) that district is still in use, the competitiveness of the district she moves to is mean-independent of $\mathrm{c}_{a_i}$, conditional on $g_i$ and $\tau$. This is:

$$\mathrm{E}[\mathrm{c}_{i\tau}|g_i,\tau,\mathrm{c}_{a_i},s_{i\tau} = 0] = \mathrm{E}[\mathrm{c}_{i\tau}|g_i,\tau,s_{i\tau} = 0] \text{ for } \tau \le \tau^l. \qquad (7)$$

By contrast, we don't make any assumptions about the relationship between $\mathrm{c}_{a_i}$ and $\mathrm{c}_{i\tau}$ in elections after a subsequent redistricting episode.

Finally, we highlight that assumptions (6) and (7) are not necessary for our empirical strategy to be valid. We make them because they permit an easier analysis of the magnitude of $\beta_\tau$ and because they are likely close to true.[8]

8. For instance, in Table A43, we present a result that is similar to assumption (6). We show that, in each $\tau$, $\mathrm{C}_{a_i\tau}$ isn't associated with a registrant's probability of being registered in her assigned district after controlling for match-groups. This reveals that $\Pr[s_{i\tau} = 1|g_i,\tau,\mathrm{c}_{a_i}] - \Pr[s_{i\tau} = 1|g_i,\tau]$ is uncorrelated with $\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]$. However, it doesn't fully prove (6), which requires that $\Pr[s_{i\tau} = 1|g_i,\tau,\mathrm{c}_{a_i}] - \Pr[s_{i\tau} = 1|g_i,\tau]$ always be zero.

### A4.3.3 Deriving the claims

We now derive the claims from Appendix A4.3.

The first claim concerns the case where no registrants have moved since the baseline election and where the assigned districts are still in use. In this case, $\mathrm{C}_{i\tau} = \mathrm{C}_{a_i\tau}$. Thus, $\beta_\tau = 1$, as claimed.

The second and third claims deal with the case where the assigned districts are still in use but where some registrants have moved. To show the claims, we obtain a formula for $\beta_\tau$ in this case. We start by deriving the formulas for $\tau = 0$ and $\tau = 1$. We then provide the general formula.

When $\tau = 0$, we have: $\mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau = 0, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau = 0]$

$$\begin{aligned}
&= \mathrm{E}[\mathrm{c}_{i0}|g_i, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{c}_{i0}|g_i] \\
&= (\mathrm{E}[\mathrm{c}_{i0}|g_i, \mathrm{c}_{a_i}, s_{i0} = 1] \cdot \Pr[s_{i0} = 1|g_i, \mathrm{c}_{a_i}] + \mathrm{E}[\mathrm{c}_{i0}|g_i, \mathrm{c}_{a_i}, s_{i0} = 0] \cdot \Pr[s_{i0} = 0|g_i, \mathrm{c}_{a_i}]) \\
&- (\mathrm{E}[\mathrm{c}_{i0}|g_i, s_{i0} = 1] \cdot \Pr[s_{i0} = 1|g_i] + \mathrm{E}[\mathrm{c}_{i0}|g_i, s_{i0} = 0] \cdot \Pr[s_{i0} = 0|g_i]) \\
&= (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \Pr[s_{i0} = 1|g_i].
\end{aligned}$$

Here, the first equality is due to the definition of $\mathrm{C}_{i\tau}$ when $\tau = 0$, the second equality uses the law of total expectation, and the last equality is due to assumptions (6) and (7). Also, we have:

$$\mathrm{C}_{a_i0} - \mathrm{E}[\mathrm{C}_{a_i0}|g_i] = \mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i].$$

Substituting the resulting quantities into Equation (5), we get:

$$\beta_0 = \frac{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2 \cdot \Pr[s_{i0} = 1|g_i]]}{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2]}.$$

When $\tau = 1$, we have: $\mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau = 1, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau = 1]$

$$\begin{aligned}
&= \mathrm{E}[\mathrm{C}_{i1}|g_i, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i1}|g_i] \\
&= \mathrm{E}[\mathrm{c}_{i0}|g_i, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{c}_{i0}|g_i] + \mathrm{E}[\mathrm{c}_{i1}|g_i, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{c}_{i1}|g_i] \\
&= (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \Pr[s_{i0} = 1|g_i] + (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \Pr[s_{i1} = 1|g_i] \\
&= (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \sum_{h=0}^{1} \Pr[s_{ih} = 1|g_i].
\end{aligned}$$

Also, we know $\mathrm{C}_{a_i1} - \mathrm{E}[\mathrm{C}_{a_i1}|g_i] = (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot 2$. Again substituting into Equation (5), we get:

$$\beta_1 = \frac{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2 \cdot \frac{1}{2} \cdot \sum_{h=0}^{1} \Pr[s_{ih} = 1|g_i]]}{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2]}.$$

In general, for $\tau \leq \tau^l$,

$$\begin{aligned}
\mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau] &= (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \sum_{h=0}^{\tau} \Pr[s_{ih} = 1|g_i], \\
\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i, \tau] &= (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot (\tau + 1), \\
\text{and} \qquad \beta_\tau &= \frac{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2 \cdot \frac{1}{\tau+1} \cdot \sum_{h=0}^{\tau} \Pr[s_{ih} = 1|g_i]|\tau]}{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2|\tau]}.
\end{aligned}$$

In practice, $\Pr[s_{i\tau} = 1|g_i, \tau]$ is less than 1 for all $\tau \geq 0$. Thus, we have shown that $\beta_\tau$ is less than 1, as claimed. Further, if $\Pr[s_{i\tau} = 1|g_i, \tau]$ is decreasing in $\tau$, then $\beta_\tau$ will decline in $\tau$, as claimed.

The last claim is about an election that occurs after a subsequent redistricting episode (i.e., $\tau > \tau^l$). The definition of $\mathrm{C}_{a_i\tau}$ implies that, in this election,

$$\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i, \tau] = (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot (\tau^l + 1).$$

Also, $\mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{C}_{i\tau}|g_i, \tau]$

$$= (\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \sum_{h=0}^{\tau^l} \Pr[s_{ih} = 1|g_i] + \sum_{h=\tau^l+1}^{\tau} (\mathrm{E}[\mathrm{c}_{ih}|g_i, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{c}_{ih}|g_i]).$$

Thus,

$$\beta_\tau = \beta_{\tau^l} + \frac{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i]) \cdot \frac{1}{\tau^l+1} \cdot \sum_{h=\tau^l+1}^{\tau}(\mathrm{E}[\mathrm{c}_{ih}|g_i, \mathrm{c}_{a_i}] - \mathrm{E}[\mathrm{c}_{ih}|g_i])|\tau]}{\mathrm{E}[(\mathrm{c}_{a_i} - \mathrm{E}[\mathrm{c}_{a_i}|g_i])^2|\tau]}.$$

From this equation, we can see that $\beta_\tau$ and $\beta_{\tau^l}$ will be equal if $\mathrm{c}_{a_i}$ has no predictive power—within match-groups—for the competitiveness that registrants experience in elections after $\tau^l$. Otherwise, they will differ. Notably, if the districts drawn in the subsequent episodes are similar to the original districts, then the second term in the above equation will be positive. As a result, $\beta_\tau$ will be greater than $\beta_{\tau^l}$.

The results derived in this section match the patterns that we find for the episode-specific first-stage coefficients, which can be visualized in Figure A6. They also match the patterns for the first-stage coefficients that vary by treatment group, as seen in Table A29. By contrast, they do not entirely match the pattern for the main first-stage coefficients, presented in Table A27. This is likely because those coefficients average over episodes that differ in how long their districts last.

## A5 The exclusion restriction in the IV model

In this section, we validate the IV model from a classical (i.e., constant treatment effects) perspective. In this framework, the model must satisfy a requirement called the exclusion restriction. Namely, after controlling for fixed effects, it must be the case that the only way $\mathrm{C}_{a_i\tau}$ is associated with $\mathrm{to}_{i\tau}$ is via $\mathrm{C}_{i\tau}$.

In this appendix, we provide evidence that the exclusion restriction holds. We first use a derivation to state the restriction in an intuitive manner. We then run empirical tests.

### A5.1 Stating the exclusion restriction

Recall that we have two main IV models. Model (2) is:

$$\begin{aligned}\mathrm{to}_{i\tau} &= \alpha_\tau \cdot \mathrm{C}_{i\tau} + \alpha_{g_i\tau} + \alpha_{i\tau}\\ \mathrm{C}_{i\tau} &= \beta_\tau \cdot \mathrm{C}_{a_i\tau} + \beta_{g_i\tau} + \beta_{i\tau}.\end{aligned}$$

Model (3) is similar, but limits $\alpha_\tau$ to be a single value $\alpha$.

In both models, the exclusion restriction is:

$$\mathrm{E}[\mathrm{C}_{a_i\tau} \cdot \alpha_{i\tau}|\tau] = 0.$$

It says: in each relative election, $\mathrm{C}_{a_i\tau}$ must not co-vary with the structural error, $\alpha_{i\tau}$.[9]

We can phrase the exclusion restriction more intuitively if we write the IV models in a way that removes the fixed effects. After de-meaning the variables, Model (2) becomes:

$$\begin{aligned}\mathrm{to}_{i\tau} - \mathrm{E}[\mathrm{to}_{i\tau}|g_i,\tau] &= \alpha_\tau \cdot (\mathrm{C}_{i\tau} - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau]) + \alpha_{i\tau}\\ \mathrm{C}_{i\tau} - \mathrm{E}[\mathrm{C}_{i\tau}|g_i,\tau] &= \beta_\tau \cdot (\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]) + \beta_{i\tau}.\end{aligned}$$

Model (3) can be written similarly, but with $\alpha$ instead of $\alpha_\tau$. From this formulation, we see that the exclusion restriction is:

$$\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]) \cdot \alpha_{i\tau}|\tau] = 0. \tag{8}$$

Here, $\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]$ is the $\tau$-specific difference between $i$'s value of $\mathrm{C}_{a_i\tau}$ and the mean value in $i$'s match-group. $\alpha_{i\tau}$ is the within-match-group-by-$\tau$ component of the factors that make up $i$'s turnout in $\tau$ other than $\mathrm{C}_{i\tau}$. Thus, the exclusion restriction says: in each relative election, $\mathrm{C}_{a_i\tau}$ must not co-vary with non-$\mathrm{C}_{i\tau}$ determinants of turnout, once variables have been de-meaned by combinations of match-group and relative election.

Broadly speaking, there are four non-$\mathrm{C}_{i\tau}$ determinants of turnout. These are: (i) characteristics of the districts that $i$ lived in during pre-redistricting elections (including pre-redistricting competitiveness), (ii) characteristics of $i$'s post-redistricting districts in chambers other than the chamber of interest; (iii) characteristics other than competitiveness for $i$'s districts in the chamber of interest; and (iv) registrant-specific factors that affect $i$'s turnout regardless of the districts in which she lives. For convenience, we label (i)-(iv) $\xi_i^{\mathrm{pre}}$, $\xi_{i\tau}^{\mathrm{oth}}$, $\xi_{i\tau}^{\mathrm{int}}$, and $\xi_{i\tau}^{\mathrm{reg}}$, respectively. Condition (8) implies that we can test the exclusion restriction by examining whether $\mathrm{C}_{a_i\tau}$ is associated with $\xi_i^{\mathrm{pre}}$, $\xi_{i\tau}^{\mathrm{oth}}$, $\xi_{i\tau}^{\mathrm{int}}$, or $\xi_{i\tau}^{\mathrm{reg}}$, after de-meaning by match-groups and $\tau$.

9. The conditioning on $\tau$ is because we allow $\beta$ to vary by $\tau$. I.e., as we explain in Appendix A4, we're using a set of instruments: the interaction of $\mathrm{C}_{a_i\tau}$ and indicators for $\tau$.

### A5.2 Tests of the exclusion restriction

We run a series of empirical tests, all of which indicate that the exclusion restriction holds.

The first test is to examine whether $\mathrm{C}_{a_i\tau}$ predicts a registrant's turnout behavior in pre-redistricting elections. Let $\tilde{\tau}$ be a pre-redistricting election (i.e., $\tilde{\tau} < 0$) and let $\tau$ be a post-redistricting election (i.e., $\tau \geq 0$). We regress pre-redistricting turnout, $\mathrm{to}_{i\tilde{\tau}}$, on $\mathrm{C}_{a_i\tau}$ and match-group fixed effects, separately for each combination of $\tau$ and $\tilde{\tau}$. These regressions are:

$$\mathrm{to}_{i\tilde{\tau}} = \phi_{\tau\tilde{\tau}} \cdot \mathrm{C}_{a_i\tau} + \phi_{g_i\tau\tilde{\tau}} + \phi_{i\tau\tilde{\tau}} \quad \text{for } \tau \geq 0 \text{ and } \tilde{\tau} < 0. \tag{9}$$

By the formula for an OLS coefficient in a model with unidimensional fixed effects, $\phi_{\tau\tilde{\tau}}$ is proportional to the following quantity:

$$\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]) \cdot (\mathrm{to}_{i\tilde{\tau}} - \mathrm{E}[\mathrm{to}_{i\tilde{\tau}}|g_i,\tilde{\tau}])|\tau,\tilde{\tau}]. \tag{10}$$

Also, in pre-redistricting elections, the treatment variable, $\mathrm{C}_{i\tilde{\tau}}$, is zero, meaning that turnout is $\mathrm{to}_{i\tilde{\tau}} = \alpha_{g_i\tilde{\tau}} + \alpha_{i\tilde{\tau}}$. In turn, quantity (10) simplifies to:

$$\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]) \cdot \alpha_{i\tilde{\tau}}|\tau,\tilde{\tau}]. \tag{11}$$

This is similar to the left-hand side of the exclusion restriction, (8). The only difference is that it uses a pre-redistricting error, $\alpha_{i\tilde{\tau}}$, instead of the desired post-redistricting error, $\alpha_{i\tau}$. Thus, we can gain insight into whether the exclusion restriction holds by seeing if quantity (11) is zero. We do this by examining the magnitude and statistical significance of the coefficient estimates for $\phi_{\tau\tilde{\tau}}$.

The results are presented in Tables A30-A32. For robustness, we provide results for three forms of $\mathrm{C}_{a_i\tau}$, which are constructed using our three measures of district competitiveness. In total, we run 72 tests, one for each combination of $\tau$ and $\tilde{\tau}$ and for each form of $\mathrm{C}_{a_i\tau}$.[10] The tests yield strong evidence that quantity (11) is zero: the coefficient estimates are all small and statistically insignificant.

We next directly assess whether $\mathrm{C}_{a_i\tau}$ is associated with non-$\mathrm{C}_{i\tau}$ determinants of turnout. We test conditions of the form:

$$\mathrm{E}[(\mathrm{C}_{a_i\tau} - \mathrm{E}[\mathrm{C}_{a_i\tau}|g_i,\tau]) \cdot (\xi_{i\tau} - \mathrm{E}[\xi_{i\tau}|g_i,\tau])|\tau] = 0,$$

where $\xi_{i\tau}$ is one of the four determinants discussed previously. We implement the tests by regressing the determinants on $\mathrm{C}_{a_i\tau}$ and match-group fixed effects, separately for each $\tau$. We then evaluate the coefficients on $\mathrm{C}_{a_i\tau}$, $\phi_\tau$.

The first determinant that we consider is $\xi_i^{\mathrm{pre}}$, the district characteristics that a registrant experienced in pre-redistricting elections. We run tests for five different characteristics: district

10. We use all combinations of $\tau$ and $\tilde{\tau}$ that exist in our data. Depending on the redistricting episode, we can observe up to five elections after redistricting ($\tau = 4$) and up to seven elections before ($\tilde{\tau} = -7$). All episodes have data in $\tau = 0$. Thus, for $\tau = 0$, we can run tests for all pre-redistricting elections, $\tilde{\tau} = -1, \ldots, -7$. Results for these tests are presented in Table A30. Next, in $\tau = 1$, we have data for all episodes except the second court-ordered revision. For these episodes, we observe turnout up to six elections prior to redistricting. Thus, we can run six tests, which are shown in Table A31. Third, for $\tau = 2$, we can run five tests, as seen in Table A32. Finally, for $\tau = 3$ and $\tau = 4$, we have data only for the decennial redistricting episodes. For these episodes, we observe turnout only for three pre-redistricting elections, $\tilde{\tau} = -1, \ldots, -3$. Importantly, $\mathrm{C}_{a_i\tau}$ never has predictive power for turnout in these elections, since we match on this turnout in creating the match-groups. As a result, we know $\phi_{\tau\tilde{\tau}} = 0$ for all the tests for $\tau = 3$ and $\tau = 4$, and we don't present results for these tests in a table.

In sum, for each version of $\mathrm{C}_{a_i\tau}$, we have seven tests for $\tau = 0$, six for $\tau = 1$, five for $\tau = 2$, and three each for $\tau = 3$ and $\tau = 4$. Thus, in total, we have 72 tests.

competitiveness, district share minority, district share Democratic, race closeness, and race spending. For each characteristic, we sum the value for a registrant's districts over all chambers and all pre-redistricting elections. We then use the sums as outcome variables in the regressions. The results are presented in Tables A33-A36. They reveal that $\mathrm{C}_{a_i\tau}$ is not associated with a registrant's pre-redistricting district experiences. In only three of the 75 tests can we reject that $\phi_\tau$ is 0.

The second determinant that we consider is the registrant-specific factor, $\xi_{i\tau}^{\text{reg}}$. We cannot observe this factor; thus, we cannot explicitly test whether it is associated with $\mathrm{C}_{a_i\tau}$. However, from prior results, we can deduce that the association is likely to be small. In particular, we've already shown that $\mathrm{C}_{a_i\tau}$ has no predictive power for turnout in any pre-redistricting election. In addition, we know that $\mathrm{C}_{a_i\tau}$ is not associated with pre-redistricting district experiences. Together, these facts mean that it does not predict pre-redistricting versions of $\xi_{i\tau}^{\text{reg}}$. As such, it likely also does not predict this factor in the desired post-redistricting election.

The third determinant that we study is $\xi_{i\tau}^{\text{oth}}$, the characteristics of a registrant's post-redistricting districts in chambers other than the chamber of interest. For this factor, we run tests for the same five characteristics as in the tests for $\xi_i^{\text{pre}}$. For each characteristic, we create an outcome variable by summing the value in a registrant's districts over all chambers other than the chamber of interest and over all post-redistricting elections from zero to $\tau$. As with $\xi_i^{\text{pre}}$, we then regress these variables on $\mathrm{C}_{a_i\tau}$ and match-group fixed effects, separately for each $\tau$. The results for the tests are presented in Tables A37-A41. They show that $\mathrm{C}_{a_i\tau}$ has little association with district characteristics in other chambers. The coefficient estimates are often statistically significant; however, this is likely due to our large sample size, as they are negligible in magnitude.[11]

The last potential threat is $\xi_{i\tau}^{\text{int}}$. This determinant captures district characteristics other than competitiveness for a registrant's districts in the chamber of interest. We discuss it in detail in Appendix A9 and show that it is not an issue.

11. For instance, using our main competitiveness measure, a one unit increase in $\mathrm{C}_{a_i\tau}$ is associated with a less than 0.02 unit increase in a registrant's post-redistricting sum of competitiveness in other chambers (Table A37). We can illustrate how small this value is by comparing it with the size of the first stage from the IV model. The first stage captures the association between $\mathrm{C}_{a_i\tau}$ and the registrant's post-redistricting sum of competitiveness in the chamber of interest. As shown in Appendix A4, the first stage ranges from 0.8 to 0.9, depending on the relative election. In other words, it is more than 40 times as large as the coefficient estimates in the tests for $\xi_{i\tau}^{\text{oth}}$. Nonetheless, in a robustness check, we run a version of the IV model that explicitly accounts for potential confounding due to $\xi_{i\tau}^{\text{oth}}$. We find that this has no impact on the results.

## Table A30: Predicting pre-redistricting turnout, $\mathrm{to}_{i\tilde{\tau}}$, using $\mathrm{C}_{a_i 0}$

| | Election prior to redistricting | | | | | | |
|---|---|---|---|---|---|---|---|
| | Seven | Six | Five | Four | Three | Two | One |
| *Panel A* | | | | | | | |
| Sum of competitiveness in a registrant's assigned district, $\mathrm{C}_{a_i\tau}$: main measure | 0.247 | -0.012 | -0.137 | -0.117 | 0 | 0 | 0 |
| | (0.735) | (0.610) | (0.409) | (0.451) | - | - | - |
| *Panel B* | | | | | | | |
| Sum of competitiveness in a registrant's assigned district, $\mathrm{C}_{a_i\tau}$: alternative measure | 0.282 | 0.022 | -0.043 | -0.044 | 0 | 0 | 0 |
| | (0.645) | (0.497) | (0.353) | (0.378) | - | - | - |
| *Panel C* | | | | | | | |
| Sum of competitiveness in a registrant's assigned district, $\mathrm{C}_{a_i\tau}$: Cook measure | 0.198 | 0.196 | -0.146 | -0.202 | 0 | 0 | 0 |
| | (0.800) | (0.666) | (0.503) | (0.503) | - | - | - |
| Turnout percentage | 24.9 | 41.8 | 44.3 | 51.2 | 37.8 | 64.3 | 50.2 |
| Clusters | 83 | 175 | 187 | 187 | 338 | 338 | 338 |
| Registrants | 1,799,908 | 2,791,776 | 2,954,963 | 2,954,963 | 5,203,371 | 5,203,371 | 5,203,371 |
| Registrant-episodes | 1,886,011 | 3,326,942 | 3,633,964 | 3,633,964 | 8,769,574 | 8,769,574 | 8,769,574 |

The table presents regression results from Equation (9). Specifically, it shows coefficient estimates and standard errors for $\phi_{\tau\tilde{\tau}}$ from regressions of $\mathrm{to}_{i\tilde{\tau}}$ on $\mathrm{C}_{a_i\tau}$ and match-group fixed effects. Each cell in the table represents a different regression. Results in different columns are for regressions that use different values of $\tilde{\tau}$. Results in different rows are for regressions that use different competitiveness measures to construct $\mathrm{C}_{a_i\tau}$. All regressions are for $\tau = 0$. Coefficient estimates and standard errors are denominated in percentage points. "Turnout percentage" is the percent of observations that turned out in the given election. "Registrants" is the number of distinct registrants in the sample. "Registrant-episodes" is the number of observations. Standard errors are clustered by baseline district in the chamber of interest. All regressions for a given value of $\tilde{\tau}$ have the same values for "Turnout percentage", "Clusters", "Registrants", and "Registrant-episodes". As a result, we provide this information in a single footer for each column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## Table A31: Predicting pre-redistricting turnout, $\mathrm{to}_{i\tilde{\tau}}$, using $\mathrm{C}_{a_i 1}$

| | Election prior to redistricting | | | | | |
|---|---|---|---|---|---|---|
| | Six | Five | Four | Three | Two | One |
| *Panel A* | | | | | | |
| Sum of competitiveness in a registrant's assigned district, $\mathrm{C}_{a_i\tau}$: main measure | 0.481 | -0.221 | -0.314 | 0 | 0 | 0 |
| | (1.29) | (0.369) | (0.629) | - | - | - |
| *Panel B* | | | | | | |
| Sum of competitiveness in a registrant's assigned district, $\mathrm{C}_{a_i\tau}$: alternative measure | 0.260 | -0.233 | -0.302 | 0 | 0 | 0 |
| | (1.01) | (0.328) | (0.595) | - | - | - |
| *Panel C* | | | | | | |
| Sum of competitiveness in a registrant's assigned district, $\mathrm{C}_{a_i\tau}$: Cook measure | 0.799 | -0.367 | -0.254 | 0 | 0 | 0 |
| | (1.40) | (0.522) | (0.891) | - | - | - |
| Turnout percentage | 26.8 | 54.2 | 43.2 | 37.3 | 63.8 | 49.0 |
| Clusters | 92 | 104 | 104 | 255 | 255 | 255 |
| Registrants | 1,352,931 | 1,604,817 | 1,604,817 | 4,486,052 | 4,486,052 | 4,486,052 |
| Registrant-episodes | 1,440,931 | 1,747,953 | 1,747,953 | 6,883,563 | 6,883,563 | 6,883,563 |

The table presents results from Equation (9). Values are analogous to those in Table A30, but are from regressions for $\tau = 1$.

Table A32: Predicting pre-redistricting turnout, $\text{to}_{i\tilde{\tau}}$, using $C_{a_i 2}$

| | Election prior to redistricting | | | | |
|---|---|---|---|---|---|
| | Five | Four | Three | Two | One |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.162 | -0.791 | 0 | 0 | 0 |
| | (0.441) | (0.621) | - | - | - |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.195 | -0.855 | 0 | 0 | 0 |
| | (0.463) | (0.721) | - | - | - |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.556 | -0.878 | 0 | 0 | 0 |
| | (1.10) | (1.60) | - | - | - |
| Turnout percentage | 30.9 | 63.6 | 29.7 | 68.9 | 42.6 |
| Clusters | 12 | 12 | 163 | 163 | 163 |
| Registrants | 307,022 | 307,022 | 3,928,260 | 3,928,260 | 3,928,260 |
| Registrant-episodes | 307,022 | 307,022 | 5,442,632 | 5,442,632 | 5,442,632 |

The table presents results from Equation (9). Values are analogous to those in Table A30, but are from regressions for $\tau = 2$.

Table A33: Predicting pre-redistricting district characteristics using $C_{a_i 0}$

| | Sum in pre-redistricting elections | | | | |
|---|---|---|---|---|---|
| | Competitiveness | Share minority | Share Democratic | Closeness | Ln. spending |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.090*** | -0.012 | -0.015 | -0.036 | 0.095 |
| | (0.032) | (0.032) | (0.023) | (0.044) | (0.162) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.092*** | -0.009 | -0.013 | -0.053 | 0.078 |
| | (0.031) | (0.031) | (0.021) | (0.044) | (0.147) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.091*** | -0.005 | -0.000 | -0.041 | 0.175 |
| | (0.033) | (0.036) | (0.028) | (0.053) | (0.208) |
| Mean of outcome variable | 10.6 | 3.67 | 5.77 | 7.34 | 2.60 |
| Clusters | 338 | 338 | 338 | 338 | 338 |
| Registrants | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 |
| Registrant-episodes | 8,769,574 | 8,769,574 | 8,769,574 | 8,769,574 | 8,769,574 |

The table presents results related to the association between $\xi_i^{\text{pre}}$ and $C_{a_i\tau}$. Specifically, it presents coefficient estimates and standard errors for $\phi_\tau$ from $\tau$-specific regressions of pre-redistricting district characteristics on $C_{a_i\tau}$ and match-group fixed effects. See Appendix A5.2 for details on these regressions. Each cell in the table represents a different regression. Results in different columns are for regressions that use the listed district or race characteristic in calculating the outcome variable. Results in different rows are for regressions that use different competitiveness measures to construct $C_{a_i\tau}$. All regressions are for $\tau = 0$. Outcome variables are calculated by summing the value of the listed characteristic in the registrant's districts over all chambers and all pre-redistricting elections. The "Competitiveness" outcome is calculated using our main competitiveness measure. Standard errors are clustered by baseline district in the chamber of interest. All regressions for a given outcome variable have the same values for "Mean of outcome variable", "Clusters", "Registrants", and "Registrant-episodes". Thus, we provide this information in a single footer for each column.

## Table A34: Predicting pre-redistricting district characteristics using $C_{a_i 1}$

| | Sum in pre-redistricting elections | | | | |
|---|---|---|---|---|---|
| | Competitiveness | Share minority | Share Democratic | Closeness | Ln. spending |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.002<br>(0.002) | 0.002<br>(0.003) | 0.002<br>(0.002) | -0.001<br>(0.004) | -0.010<br>(0.011) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.001<br>(0.002) | 0.003<br>(0.003) | 0.003<br>(0.002) | -0.002<br>(0.004) | -0.015<br>(0.011) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.001<br>(0.003) | 0.002<br>(0.003) | 0.002<br>(0.002) | -0.002<br>(0.004) | -0.012<br>(0.014) |
| Mean of outcome variable | 8.83 | 3.10 | 4.96 | 6.06 | 1.55 |
| Clusters | 255 | 255 | 255 | 255 | 255 |
| Registrants | 4,486,052 | 4,486,052 | 4,486,052 | 4,486,052 | 4,486,052 |
| Registrant-episodes | 6,883,563 | 6,883,563 | 6,883,563 | 6,883,563 | 6,883,563 |

The table presents results related to the association between $\xi_i^{\text{pre}}$ and $C_{a_i\tau}$. Values are analogous to those in Table A33. However, they are from regressions for $\tau = 1$.

## Table A35: Predicting pre-redistricting district characteristics using $C_{a_i 2}$

| | Sum in pre-redistricting elections | | | | |
|---|---|---|---|---|---|
| | Competitiveness | Share minority | Share Democratic | Closeness | Ln. spending |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.002<br>(0.001) | -0.000<br>(0.001) | 0.000<br>(0.001) | 0.000<br>(0.002) | -0.001<br>(0.007) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.001<br>(0.001) | 0.001<br>(0.001) | 0.001<br>(0.001) | -0.000<br>(0.002) | -0.004<br>(0.007) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.002<br>(0.001) | -0.001<br>(0.001) | -0.001<br>(0.001) | -0.000<br>(0.003) | -0.000<br>(0.009) |
| Mean of outcome variable | 7.41 | 2.52 | 4.22 | 5.02 | 1.57 |
| Clusters | 163 | 163 | 163 | 163 | 163 |
| Registrants | 3,928,260 | 3,928,260 | 3,928,260 | 3,928,260 | 3,928,260 |
| Registrant-episodes | 5,442,632 | 5,442,632 | 5,442,632 | 5,442,632 | 5,442,632 |

The table presents results related to the association between $\xi_i^{\text{pre}}$ and $C_{a_i\tau}$. Values are analogous to those in Table A33. However, they are from regressions for $\tau = 2$.

## Table A36: Predicting pre-redistricting district characteristics using $C_{a_i 3}$

| | Sum in pre-redistricting elections | | | | |
|---|---|---|---|---|---|
| | Competitiveness | Share minority | Share Democratic | Closeness | Ln. spending |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.002 | -0.001 | 0.000 | 0.000 | -0.002 |
| | (0.001) | (0.001) | (0.001) | (0.002) | (0.007) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.001 | 0.001 | 0.001 | -0.001 | -0.004 |
| | (0.001) | (0.001) | (0.001) | (0.002) | (0.007) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.002 | -0.001 | -0.001 | -0.000 | -0.000 |
| | (0.001) | (0.001) | (0.001) | (0.003) | (0.009) |
| Mean of outcome variable | 7.11 | 2.44 | 4.09 | 4.84 | 1.39 |
| Clusters | 151 | 151 | 151 | 151 | 151 |
| Registrants | 3,839,532 | 3,839,532 | 3,839,532 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 5,135,610 | 5,135,610 | 5,135,610 | 5,135,610 | 5,135,610 |

The table presents results related to the association between $\xi_i^{\text{pre}}$ and $C_{a_i\tau}$. Values are analogous to those in Table A33. However, they are from regressions for $\tau = 3$. Regressions are not shown for $\tau = 4$ because they are the same as those for $\tau = 3$.

## Table A37: Predicting the sum of post-redistricting district competitiveness for a registrant's districts in other chambers

| | Election relative to redistricting, $\tau$ | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.004*** | 0.005*** | 0.005*** | 0.009* | 0.011* |
| | (0.001) | (0.002) | (0.002) | (0.005) | (0.006) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.004*** | 0.005** | 0.004** | 0.008 | 0.011* |
| | (0.001) | (0.002) | (0.002) | (0.005) | (0.006) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.005*** | 0.007*** | 0.007*** | 0.012** | 0.013** |
| | (0.001) | (0.002) | (0.002) | (0.005) | (0.006) |
| Mean of outcome variable | 1.56 | 3.07 | 4.59 | 6.16 | 7.81 |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results for the association between $\xi_{i\tau}^{\text{oth}}$ and $C_{a_i\tau}$. Specifically, it presents coefficient estimates and standard errors for $\phi_\tau$ from $\tau$-specific regressions of post-redistricting district characteristics on $C_{a_i\tau}$ and match-group fixed effects. Each cell in the table represents a different regression. Results in different columns are for regressions that use different relative elections, $\tau$. Results in different rows are for regressions that use different competitiveness measures to construct $C_{a_i\tau}$. In all regressions, the outcome variable is related to district competitiveness. It is the sum of our main competitiveness measure in a registrant's districts for chambers other than the chamber of interest over the elections 0 to $\tau$. Standard errors are clustered by baseline district in the chamber of interest. All regressions for a given relative election have the same values for "Mean of outcome variable", "Clusters", "Registrants", and "Registrant-episodes". As a result, we provide this information in a single footer for each column.

## Table A38: Predicting the sum of post-redistricting district share minority for a registrant's districts in other chambers

| | Election relative to redistricting, $\tau$ | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | -0.004*** (0.001) | -0.007*** (0.002) | -0.008** (0.003) | -0.016* (0.008) | -0.021** (0.009) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | -0.003*** (0.001) | -0.007*** (0.002) | -0.006* (0.003) | -0.012 (0.009) | -0.016* (0.010) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | -0.005*** (0.001) | -0.009*** (0.002) | -0.012*** (0.003) | -0.029*** (0.008) | -0.038*** (0.010) |
| Mean of outcome variable | 0.61 | 1.23 | 1.82 | 2.46 | 3.16 |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results for the association between $\xi^{\text{oth}}_{i\tau}$ and $C_{a_i\tau}$. Values are analogous to those in Table A37. However, the outcome variable is related to district share minority. It is the sum of share minority in a registrant's districts for chambers other than the chamber of interest over the elections 0 to $\tau$.

## Table A39: Predicting the sum of post-redistricting district share Democratic for a registrant's districts in other chambers

| | Election relative to redistricting, $\tau$ | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | -0.002*** (0.001) | -0.005*** (0.001) | -0.006*** (0.002) | -0.011** (0.005) | -0.017*** (0.006) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | -0.002** (0.001) | -0.004*** (0.002) | -0.003 (0.002) | -0.007 (0.006) | -0.013* (0.007) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | -0.003*** (0.001) | -0.006*** (0.002) | -0.009*** (0.002) | -0.021*** (0.006) | -0.029*** (0.007) |
| Mean of outcome variable | 0.81 | 1.65 | 2.48 | 3.26 | 3.99 |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results for the association between $\xi^{\text{oth}}_{i\tau}$ and $C_{a_i\tau}$. Values are analogous to those in Table A37. However, the outcome variable is related to district share Democratic. It is the sum of share Democratic in a registrant's districts for chambers other than the chamber of interest over the elections 0 to $\tau$.

Table A40: Predicting the sum of post-redistricting race closeness for a registrant's districts in other chambers

| | Election relative to redistricting, $\tau$ | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.007*** (0.002) | 0.009** (0.004) | 0.005 (0.004) | 0.015 (0.013) | 0.018 (0.015) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.007*** (0.002) | 0.009** (0.004) | 0.004 (0.004) | 0.016 (0.014) | 0.020 (0.016) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.010*** (0.002) | 0.011*** (0.004) | 0.014*** (0.004) | 0.042*** (0.012) | 0.050*** (0.014) |
| Mean of outcome variable | 1.26 | 2.32 | 3.35 | 4.79 | 6.19 |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results for the association between $\xi^{\text{oth}}_{i\tau}$ and $C_{a_i\tau}$. Values are analogous to those in Table A37. However, the outcome variable is related to race closeness. It is the sum of race closeness in a registrant's districts for chambers other than the chamber of interest over the elections 0 to $\tau$.

Table A41: Predicting the sum of post-redistricting race ln. spending for a registrant's districts in other chambers

| | Election relative to redistricting, $\tau$ | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| *Panel A* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: main measure | 0.011 (0.007) | 0.026* (0.015) | -0.014 (0.019) | 0.026 (0.063) | 0.085 (0.080) |
| *Panel B* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: alternative measure | 0.010 (0.007) | 0.027 (0.017) | -0.017 (0.019) | 0.026 (0.069) | 0.105 (0.086) |
| *Panel C* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$: Cook measure | 0.014* (0.008) | 0.032** (0.016) | -0.016 (0.022) | 0.141* (0.073) | 0.222** (0.089) |
| Mean of outcome variable | 0.79 | 1.22 | 0.92 | 1.95 | 3.10 |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results for the association between $\xi^{\text{oth}}_{i\tau}$ and $C_{a_i\tau}$. Values are analogous to those in Table A37. However, the outcome variable is related to race spending. It is the sum of the natural log of race spending in a registrant's districts for chambers other than the chamber of interest over the elections 0 to $\tau$.

## A6 Effects on the probability of moving

We next examine whether being assigned to a more competitive legislative district affects the probability that a registrant moves and re-registers in a different location in North Carolina.

Table A42: The effect of $C_{a_i\tau}$ on whether a registrant is registered in his/her baseline county

| | **Election relative to redistricting, $\tau$** | | | | |
|---|---|---|---|---|---|
| | Zero | One | Two | Three | Four |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.224<br>(0.207) | 0.211<br>(0.166) | 0.136<br>(0.143) | 0.095<br>(0.168) | 0.165<br>(0.194) |
| Percent still in the baseline county | 96.4 | 94.5 | 92.0 | 90.3 | 88.0 |
| Clusters | 338 | 255 | 163 | 151 | 151 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 | 3,839,532 | 3,839,532 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 | 5,135,610 | 5,135,610 |

The table presents results from Equation (12). Specifically, it displays coefficient estimates and standard errors for $\mu_\tau$ from $\tau$-specific regressions of $s_{i\tau}$ on $C_{a_i\tau}$ and match-group fixed effects. The outcome variable is an indicator for whether the registrant is still registered in his or her baseline county in election $\tau$. Coefficient estimates and standard errors are denominated in percentage points. "Percent still in the baseline county" is the percent of baseline registrants who are still registered in their baseline county in the listed election. It is the mean of the outcome variable in the election. Standard errors are clustered by baseline district in the chamber of interest. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

This question is interesting for two reasons. First, it lends insight into the mechanisms driving the first-stage coefficients in the IV model. Second, it provides evidence on whether being assigned to a more competitive district affects the probability of moving and re-registering in a different state. The latter question is important because we do not observe registration and turnout in states other than North Carolina. Instead, our outcome variables capture only whether individuals who were registered in the baseline election turn out in North Carolina. If competitiveness influences the probability of moving to a different state, then the causal effects we recover may be biased by selective attrition.

Table A43: The effect of $C_{a_i\tau}$ on whether a registrant is registered in his/her assigned district

| | **Election rel. to redistricting, $\tau$** | | |
|---|---|---|---|
| | Zero | One | Two |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | 0.185<br>(0.618) | 0.631<br>(0.493) | 0.530<br>(0.448) |
| Percent still in the assigned district | 92.2 | 70.5 | 62.4 |
| Clusters | 338 | 255 | 163 |
| Registrants | 5,203,371 | 4,486,052 | 3,928,260 |
| Registrant-episodes | 8,769,574 | 6,883,563 | 5,442,632 |

The table presents results similar to those in Table A42 but for a different outcome variable. The outcome is an indicator for whether the registrant is still registered in his or her assigned district in election $\tau$. This variable is set to 0 if the assigned district no longer exists in $\tau$. Results are omitted for relative elections $\tau = 3$ and $\tau = 4$. This is because we know $\mu_\tau = 0$ in these elections. Namely, all districts are modified by subsequent redistricting episodes prior to $\tau = 3$; thus, $s_{i\tau} = 0$ for all registrants in elections $\tau \geq 3$. Coefficient estimates and standard errors are denominated in percentage points. "Percent still in the assigned district" is the percent of baseline registrants who are still registered in their assigned district in the listed election. It is the mean of the outcome variable in the election. Standard errors are clustered by baseline district in the chamber of interest.

In the analysis, we consider two variables related to moving. The first is an indicator for whether, in election $\tau$, a registrant is still registered in his or her baseline county. The second is an indicator for whether the registrant is still registered in his or her assigned district, $a_i$. This latter variable is set to 0 if the assigned district is no longer in use in $\tau$.

We recover effects by running $\tau$-specific regressions of each variable on the instrument, $\mathrm{C}_{a_i\tau}$, and match-group fixed effects. Specifically, for each $\tau$, we fit the model:

$$s_{i\tau} = \mu_\tau \cdot \mathrm{C}_{a_i\tau} + \mu_{g_i\tau} + \mu_{i\tau}. \tag{12}$$

In this equation, $s_{i\tau}$ is one of the variables related to moving, and $\mu_\tau$ is the coefficient of interest.

The results from the regressions are presented in Tables A42 and A43. Table A42 is for whether a registrant is still registered in his or her baseline county; Table A43 is for whether the registrant is still registered in his or her assigned district.

The results reveal that being assigned to a more competitive district has no effect on the probability of moving within North Carolina. In all regressions, the coefficient estimate for $\mu_\tau$ is small and statistically insignificant. This implies that the first stage is not impacted by differences in the probability of moving out of the assigned district. In addition, it is suggestive evidence that there is no effect of assigned competitiveness on the probability of moving to a different state.

## A7 The IV model with treatment effect heterogeneity

In a constant treatment effects framework, validity of the IV model requires only that the model have a non-zero first stage and satisfy the exclusion restriction. In a framework with treatment effect heterogeneity, the model must instead satisfy four requirements: a non-zero first stage, independence, exclusion, and monotonicity. Further, the causal effect that the model recovers has a special interpretation. We now discuss the IV model in this more general framework. The discussion applies equivalently to Models (2) and (3).

The independence and exclusion requirements are related to the exclusion restriction from the constant treatment effects framework. Loosely speaking, independence is that $\mathrm{C}_{a_i\tau}$ is as good as randomly assigned within match-groups. More precisely, it is that, in each relative election $\tau$, $\mathrm{C}_{a_i\tau}$ is independent of $i$'s potential outcomes and potential treatment statuses, once we control for match-groups. In our setting, potential outcomes capture whether $i$ would turn out under alternative values of the treatment, $\mathrm{C}_{i\tau}$. Potential treatment statuses are the values of $\mathrm{C}_{i\tau}$ that $i$ would obtain under alternative values of $\mathrm{C}_{a_i\tau}$. In Appendix A5, we showed that $\mathrm{C}_{a_i\tau}$ is not related to major sources of concern, including pre-redistricting experiences, experiences in chambers other than the chamber of interest, and registrant factors that affect $i$'s turnout regardless of her districts. As such, we believe that the independence requirement is satisfied in our setting.

Exclusion is that $\mathrm{C}_{a_i\tau}$ affects turnout only via the desired causal channel. In our case, we are interested in the treatment effect of living in a more or less competitive district. Thus, the desired causal channel is that $\mathrm{C}_{a_i\tau}$ affects $\mathrm{C}_{i\tau}$ which in turn affects turnout. Note that there are a number of sub-channels that form part of this broader channel. As we explain in the conceptual overview in the main text, living in a more or less competitive district could affect turnout via exposure to close or non-close races, responses by candidates and campaigns, the development of social norms, etc. Thus, $\mathrm{C}_{a_i\tau}$ may also affect these other variables. This is not a violation of exclusion so long as the relationship between $\mathrm{C}_{a_i\tau}$ and the other variables is due to $\mathrm{C}_{i\tau}$.

One way in which exclusion could be violated in our setting is through the existence of majority-minority districts. As we explain in Appendix A9, these generate a negative correlation between a district's competitiveness and its share minority and share Democratic. Thus, being assigned to a less competitive district tends to cause a registrant to experience districts that are more heavily minority and Democratic. If a district's racial and partisan composition matter for turnout, then assigned competitiveness may affect turnout in part through these race and party channels, which are distinct from the competitiveness channel.[12] Fortunately, we are able to assess the extent of this bias in Appendix A9, and we find that it is minimal.

Monotonicity is the assumption that $\mathrm{C}_{a_i\tau}$ weakly increases $\mathrm{C}_{i\tau}$ for all $i$. This assumption can be understood in relation to compliance types. First, it says that there can be compliers, defined as people who experience additional competitiveness when assigned to a more competitive district and less competitiveness when assigned to a less competitive one. Second, there can also be people for whom assigned competitiveness does not matter for experienced competitiveness. By contrast, there can be no defiers, defined as people for whom $\mathrm{C}_{a_i\tau}$ reduces $\mathrm{C}_{i\tau}$. (These people experience less competitiveness when assigned to a more competitive district and more competitiveness when assigned to a less competitive one.)

The monotonicity assumption is not testable; however, we believe that it holds at least approx-

12. The reason the two sets of channels are distinct is that there are places where competitiveness is not correlated with share minority and share Democratic (such as states with low minority shares). As such, being heavily minority and Democratic is not a core feature of uncompetitive districts; instead, it's merely an attribute that North Carolina's uncompetitive districts often have.

imately in our setting. To see this, it's useful to think carefully about defiers. Defiance could arise either due to people's migration decisions or due to the decisions of policymakers in subsequent redistricting episodes. In the first case, defiers are registrants who move between districts in a way that counteracts the effect of assigned competitiveness. Specifically, their migration decisions mean that they experience less (more) competitiveness when assigned to a more (less) competitive district.[13] In the second case, defiers are registrants whose subsequent district placements undo their assigned competitiveness.[14] It's likely that there are few—if any—people in these groups. As evidence, if the first story were to hold widely, then we may expect people to be more likely to move if assigned to certain types of districts (e.g., uncompetitive ones). However, in Appendix A6, we show that assigned competitiveness does not predict the probability of moving from one's district. Also, if the second story were to hold widely, then we would expect there to be a negative correlation between assigned competitiveness and the competitiveness that registrants experience in elections after subsequent redistricting episodes. However, in Appendix A4, we show that this correlation is either zero or slightly positive, depending on the episode. In sum, we can't rule out that there are defiers. Nonetheless, it seems plausible that their number is small and that any bias they generate is modest.

Interestingly, in our setting, we believe that the number of people for whom $\mathrm{C}_{i\tau}$ is unaffected by $\mathrm{C}_{a_i\tau}$ is also small. The main example of such a person is someone who moves out of their assigned district in the time between redistricting and the first post-redistricting election.[15] Table A43 shows that less than 8% of registrants fall into this category. Thus, in our setting, most registrants are compliers—although registrants may differ in their degree of compliance.

Finally, IV has a special interpretation under treatment effect heterogeneity. The causal effect that it recovers is a local average treatment effect (LATE), defined as an average treatment effect for compliers. In other words, the LATE is for people who are induced to experience more (less) competitiveness by being assigned to a more (less) competitive district. It is not for people whose experienced competitiveness is unaffected by assigned competitiveness. In addition, it is not for defiers, who are assumed to not exist.

Despite its limited nature, the LATE is interesting, for at least three reasons. First, since most registrants in our setting are compliers, it is relevant for most registrants. Second, we can't think of a story where non-compliers are substantially different from compliers.[16] Thus, effects for compliers and non-compliers may be similar, at least once we subset by chamber type and by observable registrant characteristics. Third, compliers are the registrants who are influenced by assigned competitiveness, which is what policymakers control in redistricting. Thus, compliers are a key group from a policy perspective.

13. An example is a person who is on the margin of moving to a competitive district. The person decides to move if assigned to a highly uncompetitive district but not if assigned to a moderately uncompetitive one.

14. If these people are assigned to a more (less) competitive district in the initial redistricting episode, then they get put into less (more) competitive districts in subsequent episodes. Further, the subsequent districts are so much less (more) competitive that the people experience lower (higher) cumulative competitiveness, $\mathrm{C}_{i\tau}$, than if they had initially been placed in a less (more) competitive district. An example where this would occur is if policymakers decide that they want to balance out gaps in $\mathrm{C}_{i\tau}$.

15. A person who moves later is likely to be a complier, as the person's value of $\mathrm{C}_{i\tau}$ will be affected by the person's experiences in the elections in which he or she remained in the assigned district.

16. As mentioned, the major difference between compliers and non-compliers is when the groups move out of their districts. Non-compliers leave their districts in the time between redistricting and the first post-redistricting election. By contrast, compliers leave after the first post-redistricting election—or not at all.

## A8 Turnout's dependence on cumulative competitiveness

In the main text, we provided evidence that turnout depends on cumulative competitiveness. We did this in two ways. First, in Figures 3 and A1, we showed visually that the effects of assigned competitiveness on turnout track those on $C_{i\tau}$, the sum of competitiveness in a registrant's districts in the chamber of interest in elections since redistricting. Second, in Table 1, we showed that estimates of $\alpha_\tau$, the coefficients on $C_{i\tau}$ in Model (2), are stable across relative elections. In this appendix, we present further evidence that turnout depends on cumulative competitiveness. We do this by directly estimating the impacts of current and lagged competitiveness.

Turnout depending on cumulative competitiveness means that the impacts of current and lagged competitiveness are the same. An alternative possibility is that turnout is a moving-average process. This is where the impacts of lagged competitiveness decline as the number of lags increases. The redistricting natural experiment allows us to distinguish between these possibilities, but only for a number of lags equal to one less than the number of post-redistricting elections (4). By contrast, it does not permit us to comment on the impacts of additional lags of competitiveness (> 4). This is because it does not generate variation in the competitiveness that registrants experience in pre-redistricting elections. (In Figure 3, we found that there is no difference—within match-groups—in pre-redistricting competitiveness for registrants placed into more versus less competitive districts during redistricting.) As such, we can exploit lagged competitiveness that occurs in post-redistricting elections, but not pre-redistricting elections.

To understand our strategy, write turnout in relative election $\tau$ as a function of current competitiveness ($c_{i\tau}$), lagged competitiveness for the number of lagged elections between $\tau$ and redistricting ($c_{i\tau-l}$ for $l = 1, \ldots, \tau$), a match-group-by-$\tau$ fixed effect ($\kappa_{g_i\tau}$), and an error term ($\kappa_{i\tau}$). It is:

$$\text{to}_{i\tau} = \sum_{l=0}^{\tau} \kappa_l \cdot c_{i\tau-l} + \kappa_{g_i\tau} + \kappa_{i\tau}, \tag{13}$$

for $\tau = 0, \ldots, 4$. In the equation, $\kappa_l$ is the effect of a one-unit increase in the $l^{\text{th}}$ lag of competitiveness. The impacts of all lags beyond the $\tau^{\text{th}}$ are captured in the $\kappa_{g_i\tau}$ fixed effect and the $\kappa_{i\tau}$ error. If turnout depends on cumulative competitiveness—at least over the number of lags that we can identify—then all $\kappa_l$ coefficients should equal $\alpha$, the coefficient on $C_{i\tau}$ in Model (3). If turnout is instead a moving-average process, then $\kappa_l$ should decline as $l$ increases.

We recover the $\kappa_l$ coefficients by using an IV model. Specifically, we instrument for $c_{i\tau-l}$ using the interaction of assigned competitiveness, $c_{a_i}$, treatment group, and relative election. Identification comes from the fact that, after controlling for match-groups, assigned competitiveness affects different combinations of current and lagged competitiveness for different treatment groups and relative elections. Thus, we can trace how differences in turnout align with these instrument-induced differences in current and lagged competitiveness. The exclusion restriction is that the instruments are not associated with $\kappa_{i\tau}$. This would be violated if they have predictive power for lags of competitiveness that occur in pre-redistricting elections (i.e., $c_{i\tau-l}$ for $l > \tau$); however, we have shown that this is not the case.

The identifying variation for the IV model is summarized in Table A44. The table presents results from the first-stage regressions, which are regressions of $c_{i\tau-l}$ on the instruments and match-group-by-$\tau$ fixed effects. The table reveals which treatment groups and relative elections contribute to identifying which $\kappa_l$ coefficients.

In relative election $\tau = 0$, assigned competitiveness, $c_{a_i}$, has predictive power only for current competitiveness, $c_{i\tau}$. This is the case for all treatment groups and is because $\tau = 0$ is the first election after redistricting. Thus, $\tau = 0$ helps to identify only $\kappa_0$, the coefficient on current competitiveness. In $\tau = 1$, $c_{a_i}$ is associated with $c_{i\tau}$ and $c_{i\tau-1}$ for all treatment groups. However, the association with $c_{i\tau}$ for Group C is half as large. This is because, by $\tau = 1$, there has already been a subsequent redistricting episode for this group. Thus, $\tau = 1$ helps to identify both $\kappa_0$ and $\kappa_1$, but the extent to which it contributes to $\kappa_0$ varies by group. In $\tau = 2$, $c_{a_i}$ is associated with $c_{i\tau}$, $c_{i\tau-1}$, and $c_{i\tau-2}$ for Group A and with $c_{i\tau-1}$ and $c_{i\tau-2}$ for Group B. The lack of association with $c_{i\tau}$ for Group B is again because there has now been a subsequent redistricting episode for this group. Also, Group C doesn't help to identify any coefficients in $\tau = 2$; this is because its episodes have no data after $\tau = 1$. Next, in $\tau = 3$, $c_{a_i}$ is associated with $c_{i\tau}$, $c_{i\tau-1}$, $c_{i\tau-2}$, and $c_{i\tau-3}$ for Group A and with $c_{i\tau-2}$ and $c_{i\tau-3}$ for Group B. However, the association with $c_{i\tau}$ for Group A is weak. Finally, in $\tau = 4$, for Group A, $c_{a_i}$ is weakly associated with $c_{i\tau}$ and $c_{i\tau-1}$ and strongly associated with $c_{i\tau-2}$, $c_{i\tau-3}$, and $c_{i\tau-4}$. For Group B, it is associated with $c_{i\tau-3}$ and $c_{i\tau-4}$.

Table A44: The first stage in Model (13)

| | $c_{i\tau}$ | $c_{i\tau-1}$ | $c_{i\tau-2}$ | $c_{i\tau-3}$ | $c_{i\tau-4}$ |
|---|---|---|---|---|---|
| Group A: $\tau = 0$ | 0.869*** | | | | |
| | (0.006) | | | | |
| Group A: $\tau = 1$ | 0.822*** | 0.869*** | | | |
| | (0.007) | (0.006) | | | |
| Group A: $\tau = 2$ | 0.752*** | 0.822*** | 0.869*** | | |
| | (0.008) | (0.007) | (0.006) | | |
| Group A: $\tau = 3$ | 0.175*** | 0.752*** | 0.822*** | 0.869*** | |
| | (0.045) | (0.008) | (0.007) | (0.006) | |
| Group A: $\tau = 4$ | 0.076*** | 0.175*** | 0.752*** | 0.822*** | 0.869*** |
| | (0.022) | (0.045) | (0.008) | (0.007) | (0.006) |
| Group B: $\tau = 0$ | 0.875*** | | | | |
| | (0.008) | | | | |
| Group B: $\tau = 1$ | 0.827*** | 0.875*** | | | |
| | (0.010) | (0.008) | | | |
| Group B: $\tau = 2$ | 0.003 | 0.827*** | 0.875*** | | |
| | (0.010) | (0.010) | (0.008) | | |
| Group B: $\tau = 3$ | -0.005 | -0.006 | 0.837*** | 0.879*** | |
| | (0.005) | (0.005) | (0.011) | (0.009) | |
| Group B: $\tau = 4$ | -0.022 | -0.005 | -0.006 | 0.837*** | 0.879*** |
| | (0.015) | (0.005) | (0.005) | (0.011) | (0.009) |
| Group C: $\tau = 0$ | 0.874*** | | | | |
| | (0.006) | | | | |
| Group C: $\tau = 1$ | 0.546*** | 0.898*** | | | |
| | (0.068) | (0.005) | | | |
| Clusters | 338 | 338 | 338 | 338 | 338 |
| Registrants | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 |
| Registrant-episode-elections | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 |

The table presents results from the first-stage regressions associated with Model (13). Each column is for a regression of the listed competitiveness variable on the instruments and match-group-by-$\tau$ fixed effects. The regressions use data from all post-redistricting elections. Lagged competitiveness is set to 0 if it occurs in a pre-redistricting election. The instruments are the interaction of $c_{a_i}$, treatment group, and relative election. They are labeled based on the treatment group and relative election. Group A is the redistricting episodes in which districts last for three elections. Group B (C) is the episodes in which districts last for two elections (one election). Standard errors are clustered by baseline district in the chamber of interest. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results from Model (13) are displayed in Table A45. They support the claim that turnout depends on cumulative competitiveness. While there is some variation in the coefficient estimates for $\kappa_l$, there is no trend across $l$. In addition, most of the variation seems to be noise. In a joint test, we cannot reject that all $\kappa_l$ coefficients are equal to 1.30, the estimate of $\alpha$ in Table 1.

Table A45: The effects of current and lagged competitiveness

| | (1) |
|---|---|
| Current competitiveness, $c_{i\tau}$ | $1.77^{***}$ |
| | (0.471) |
| First lag of competitiveness, $c_{i\tau-1}$ | 0.311 |
| | (0.496) |
| Second lag of competitiveness, $c_{i\tau-2}$ | $1.91^{***}$ |
| | (0.671) |
| Third lag of competitiveness, $c_{i\tau-3}$ | $1.08^{*}$ |
| | (0.630) |
| Fourth lag of competitiveness, $c_{i\tau-4}$ | $1.67^{***}$ |
| | (0.619) |
| Turnout percentage | 58.1 |
| F-statistic | 1.03 |
| p-value | 0.401 |
| Clusters | 338 |
| Registrants | 5,203,371 |
| Registrant-episode-elections | 31,366,989 |

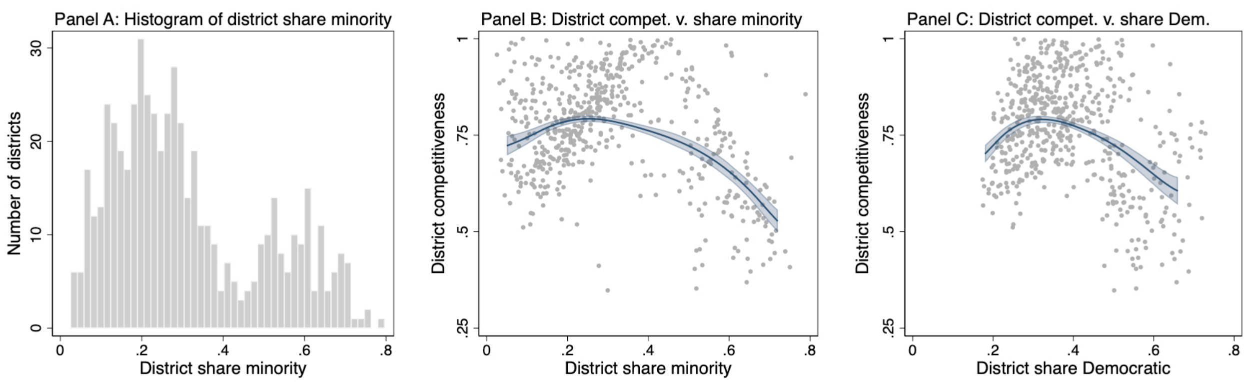
The table displays results from Model (13). Specifically, it shows coefficient estimates and standard errors for $\kappa_l$ in a two-stage least squares regression of $to_{i\tau}$ on current and lagged competitiveness and match-group-by-$\tau$ fixed effects. “F-statistic” and “p-value” are from a test that all $\kappa_l$ coefficients are equal to 1.30. See the notes to Table A44 for details on the first-stage regressions, the sample, and clustering.

In sum, we have provided considerable evidence that turnout depends on cumulative competitiveness, at least for a certain length of time. In Figure 3 (Figure A1), we showed that treatment effects on turnout do not decay for at least two (three) elections after differences in current competitiveness disappear. This suggests that lagged competitiveness has the same impact as current competitiveness for at least (two) three lags. In Table 1, we showed that estimates of $\alpha_\tau$ in Model (2) are stable for at least five elections. This suggests that lagged and current competitiveness have the same impact for at least four lags. Finally, in this appendix, we directly estimated the impacts of current and lagged competitiveness. We again found that these are similar for at least four lags.

# A9 The roles of other district attributes

A potential issue with our empirical strategy is that a district's competitiveness may be correlated with its other attributes. If so, then people who live in more competitive districts will be subject to a bundle of treatments. The IV model will not isolate the causal effect of competitiveness alone. Instead, it will recover the combined effect of competitiveness and of the other attributes that are associated with competitiveness.

Figure A7: The relationship between competitiveness, share minority, and share Democratic




The figure plots information on the relationship between district competitiveness, district share minority, and district share Democratic. The sample is all districts that were used in North Carolina during the 2012 through 2020 elections.

This concern is relevant in North Carolina because of the existence of "majority-minority" districts. As seen in Figure A7, a substantial fraction of the state's districts have a large share of registrants who are racial minorities (Panel A). These districts tend to be highly Democratic and highly uncompetitive. As a result, there is an overall negative correlation between a district's competitiveness and both its share minority (Panel B) and its share Democratic (Panel C).[17] Prior research finds that a district's racial and partisan makeup each affect voter turnout. In particular, the research finds that registrants are more likely to turn out when they live in districts with a larger share of people who are of either their same race (Fraga 2016) or party (Fraga, Moskowitz, and Schneer 2021). As such, in our setting, the effect of living in a competitive district may reflect these other treatments.

The issue of correlated treatments is mitigated by the fact that the race and party channels are "match effects". That is, they have different signs for different types of registrants. As an example, the minority share is expected to increase turnout for minority registrants but decrease turnout for white registrants. Similarly, the share Democratic should raise turnout for Democrats but reduce turnout for other registrants. Over all registrants, the match effects should mostly cancel; thus, the influence of the race and party channels should be small. On the other hand, the channels may have an appreciable impact when we calculate effects for groups of registrants that are homogenous in terms of race or party.

17. Table A46 shows that the negative correlations remain when we control for match-groups. Specifically, it reveals that the instrument, $C_{a_i\tau}$, predicts both the share minority and the share Democratic in a registrant's districts, conditional on match-group-by-$\tau$ fixed effects.

### A9.1 Strategy

We deal with the issue of correlated treatments by presenting two sets of results. In our main results, we do not attempt to adjust for the race or party channels. These results reveal the total effect of living in a more or less competitive district in North Carolina, inclusive of the fact that uncompetitive districts in the state are often heavily minority and Democratic. The total effects are relevant in North Carolina and in other states with majority-minority districts. In a second set of results, we isolate the partial effect of competitiveness, accounting for the impacts of race and party. The partial effects are relevant in settings where a district's competitiveness is not associated with its racial or partisan composition.

We obtain the second set of results by fitting the IV model on a trimmed sample. Namely, we exclude registrants who are assigned to districts with an extremely high minority share. This kills the associations between the instrument, $\mathrm{C}_{a_i\tau}$, and the values of district share minority or share Democratic that registrants experience.[18] Thus, on the trimmed sample, the instrument is related to turnout only via its association with district competitiveness. In turn, the IV model estimated on this sample identifies the partial effect of competitiveness.

In practice, we find that the partial and total effects are similar. Thus, in the main paper, we focus only on the total effects, which are more precise.

### A9.2 Results

Table A48 presents IV results for the full and trimmed samples. The values are coefficient estimates and standard errors for $\alpha$ in Model (3). Panel A is for the full sample (the "total effects"), while Panel B is for the trimmed sample (the "partial effects").

The table provides two takeaways. First, as predicted, coefficient estimates are similar for the two samples when using registrants of all races and parties (the "All" columns). The full-sample estimate of $\alpha$ is 1.30, as in Table 1. The trimmed-sample estimate is 1.48.

Second, total and partial effects are also similar when subsetting by race and party. For each race-party group, the difference between the two effects is modest and statistically insignificant. That said, the signs of the differences align with what would be predicted based on prior research. For instance, for minorities, we would expect the partial effect to be larger than the total effect. This is because minorities gain a positive racial and partisan match when they live in uncompetitive districts. In line with this, Table A48 shows that the partial effect is larger for minorities, but only to a limited degree. Next, for white-Republicans and white-Unaffiliated registrants, the partial effect should be smaller than the total effect. This is because these groups receive negative match effects from uncompetitive districts. We again find the expected result; however, the differences are again limited. Finally, for white-Democrats, the total and partial effects should be similar. This is because white-Democrats gain conflicting match effects from uncompetitive districts: they receive a positive partisan match and a negative racial match. As expected, the effects are similar; however, the partial effect is slightly larger. In sum, the race and party heterogeneity is consistent with the existence of race- and party-based match effects. Yet, the match effects are too small to cause significant distortion.

18. We trim the sample separately by a registrant's race and party. This way, we can kill the associations both overall and conditional on a registrant's type. For racial minorities and white-Democrats, we drop registrants who are assigned to districts that are more than 61.5% minority. For white-Unaffiliated registrants and white-Republicans, the cutoffs are 62.5% and 63.5%. Table A47 presents results from regressions of share minority and share Democratic on $\mathrm{C}_{a_i\tau}$ and match-group-by-$\tau$ fixed effects for the trimmed sample. It shows that the trimming is successful: the coefficients on $\mathrm{C}_{a_i\tau}$ are zero, both overall and by registrant type.

Table A46: The association between $C_{a_i\tau}$ and the sums of post-redistricting share minority and share Democratic in the chamber of interest: full sample

| | All | White Dem. | White Rep. | White Unaffil. | Minority |
|---|---|---|---|---|---|
| *Panel A: District share minority* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | -0.582*** | -0.549*** | -0.492*** | -0.507*** | -0.712*** |
| | (0.071) | (0.084) | (0.084) | (0.071) | (0.062) |
| Mean of outcome variable | 0.85 | 0.78 | 0.71 | 0.73 | 1.17 |
| Clusters | 338 | 338 | 337 | 336 | 336 |
| Registrants | 5,203,371 | 1,057,919 | 1,560,055 | 1,222,064 | 1,500,825 |
| Registrant-episode-years | 31,366,989 | 6,730,584 | 9,490,452 | 6,528,673 | 8,617,280 |
| *Panel B: District share Democratic* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | -0.319*** | -0.310*** | -0.240*** | -0.278*** | -0.404*** |
| | (0.044) | (0.046) | (0.061) | (0.043) | (0.039) |
| Mean of outcome variable | 1.11 | 1.11 | 1.00 | 1.02 | 1.29 |
| Clusters | 338 | 338 | 337 | 336 | 336 |
| Registrants | 5,203,371 | 1,057,919 | 1,560,055 | 1,222,064 | 1,500,825 |
| Registrant-episode-years | 31,366,989 | 6,730,584 | 9,490,452 | 6,528,673 | 8,617,280 |

The table presents results for regressions of post-redistricting district characteristics on $C_{a_i\tau}$ and match-group-by-$\tau$ fixed effects. The outcome variable in Panel A is related to district share minority. It is the sum of share minority in a registrant's districts for the chamber of interest over the elections 0 to $\tau$. The outcome in Panel B is an analogous sum but for district share Democratic. Results in different columns are for regressions that use the listed sets of registrants. Standard errors are clustered by baseline district in the chamber of interest. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A47: The association between $C_{a_i\tau}$ and the sums of post-redistricting share minority and share Democratic in the chamber of interest: trimmed sample

| | All | White Dem. | White Rep. | White Unaffil. | Minority |
|---|---|---|---|---|---|
| *Panel A: District share minority* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | -0.060 | -0.067 | -0.034 | -0.098 | -0.035 |
| | (0.081) | (0.073) | (0.101) | (0.080) | (0.101) |
| Mean of outcome variable | 0.68 | 0.65 | 0.64 | 0.64 | 0.84 |
| Clusters | 306 | 294 | 306 | 299 | 290 |
| Registrants | 4,428,467 | 919,260 | 1,458,917 | 1,126,341 | 1,044,539 |
| Registrant-episode-years | 23,975,039 | 5,352,542 | 8,440,837 | 5,626,651 | 4,555,009 |
| *Panel B: District share Democratic* | | | | | |
| Sum of competitiveness in a registrant's assigned district, $C_{a_i\tau}$ | -0.016 | -0.045 | 0.048 | -0.050 | -0.025 |
| | (0.074) | (0.066) | (0.092) | (0.073) | (0.080) |
| Mean of outcome variable | 0.99 | 1.02 | 0.96 | 0.96 | 1.07 |
| Clusters | 306 | 294 | 306 | 299 | 290 |
| Registrants | 4,428,467 | 919,260 | 1,458,917 | 1,126,341 | 1,044,539 |
| Registrant-episode-years | 23,975,039 | 5,352,542 | 8,440,837 | 5,626,651 | 4,555,009 |

The table is analogous to Table A46. However, the sample excludes registrants who are assigned to districts with large minority shares. For racial minorities, white-Democrats, white-Unaffiliated registrants, and white-Republicans, it drops registrants assigned to districts with shares greater than 0.615, 0.615, 0.625, and 0.635, respectively.

Table A48: The effects of district competitiveness for the full and trimmed samples

| | All | White: Dem. | White: Rep. | White: Unaffil. | Minority |
|---|---|---|---|---|---|
| *Panel A: Full sample ("total effects")* | | | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 2.10*** | 1.30*** | 1.73*** | 0.361 |
| | (0.234) | (0.284) | (0.293) | (0.277) | (0.385) |
| Turnout percentage | 58.1 | 62.0 | 63.7 | 51.8 | 53.4 |
| Clusters | 338 | 338 | 337 | 336 | 336 |
| Registrants | 5,203,371 | 1,057,919 | 1,560,055 | 1,222,064 | 1,500,825 |
| Registrant-episode-elections | 31,366,989 | 6,730,584 | 9,490,452 | 6,528,673 | 8,617,280 |
| *Panel B: Trimmed sample ("partial effects")* | | | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.48*** | 2.20*** | 1.03** | 1.69*** | 0.588 |
| | (0.336) | (0.395) | (0.449) | (0.384) | (0.792) |
| Turnout percentage | 59.7 | 63.0 | 64.7 | 53.1 | 54.6 |
| Clusters | 306 | 294 | 306 | 299 | 290 |
| Registrants | 4,428,467 | 919,260 | 1,458,917 | 1,126,341 | 1,044,539 |
| Registrant-episode-elections | 23,975,039 | 5,352,542 | 8,440,837 | 5,626,651 | 4,555,009 |

The table shows how the effects of district competitiveness are influenced by the correlations between competitiveness, share minority, and share Democratic. Specifically, it presents results from Model (3) for two samples. Panel A uses the full sample. Panel B uses a trimmed sample that excludes registrants assigned to districts with high minority shares; see Appendix A9.1 for details. Results in different columns are for regressions that use the listed sets of registrants. The "All" column in Panel A matches the "All" column in Table 1. The "White-Dem.", "White-Rep.", "White-Unaffil.", and "Minority" columns in Panel A correspond with the results by race and party that are presented in Figure 4. All other details are the same as in Table 1.

## A10 Robustness to approaches that do not involve stacking

One complication with our empirical strategy is that it relies on stacking. In an attempt to gain statistical power, we combine data from multiple redistricting episodes, repeating a registrant's observations each time the registrant is registered in an episode's baseline election. It is conceivable that this approach could introduce unknown distortions.

In this appendix, we explore the role of stacking by re-running analyses on samples that do not involve stacking. We do this in two steps. First, in Table A49, we run the main IV model, Model (3), separately for each of the decennial redistricting episodes. These are the episodes for the U.S. House, the NC Senate, and the NC House that occurred in advance of the 2012 election. By running the model on the decennial episodes one at a time, we ensure that the samples do not repeat observations for the same registrant. Instead, in each sample, the number of registrants is the same as the number of registrant-episode combinations. As a comparison, we also run the model on samples that stack all the episodes for a given chamber. By comparing results for the two sets of samples, we can gain insight into the influence of stacking.

Table A49 suggests that stacking is not distortionary. The results for the samples that do not involve stacking (Panel B) are quite similar to those for the samples that do (Panel A).

Table A49: The turnout effects of district competitiveness by chamber: with and without stacking

| | **U.S. House** | **NC Senate** | **NC House** |
|---|---|---|---|
| *Panel A: Full sample* | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 2.03*** | 0.921* | 1.33*** |
| | (0.738) | (0.530) | (0.256) |
| Turnout percentage | 58.3 | 58.3 | 57.8 |
| Clusters | 36 | 74 | 228 |
| Registrants | 1,586,751 | 1,679,034 | 4,025,851 |
| Registrant-episodes | 1,700,565 | 1,910,642 | 5,158,367 |
| Registrant-episode-elections | 6,505,033 | 7,014,716 | 17,847,240 |
| *Panel B: Only the decennial episodes* | | | |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 2.29** | 0.904 | 1.35*** |
| | (0.776) | (0.559) | (0.265) |
| Turnout percentage | 56.6 | 56.7 | 56.3 |
| Clusters | 13 | 32 | 106 |
| Registrants | 1,047,606 | 1,193,592 | 2,894,412 |
| Registrant-episodes | 1,047,606 | 1,193,592 | 2,894,412 |
| Registrant-episode-elections | 5,238,030 | 5,967,960 | 14,472,060 |

The table displays results from Model (3) by legislative chamber. Panel A uses all the episodes for a given chamber and thus relies on stacking. Panel B uses just the decennial episode for the chamber and thus does not rely on stacking. All other details are the same as in Table 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We next conduct a similar analysis but for the empirical strategy that was used in Table 5. Recall that this strategy defines regions based only on pre-redistricting districts and thus allows people in match-groups to differ in assigned districts for multiple chambers. The strategy was useful in that it let us see how the effects of district competitiveness aggregate across chambers. A key feature of it was that we had to change the definition of a redistricting episode. Instead of being an instance where districts were redrawn for a particular chamber, the definition became an instance where districts were redrawn for any chamber. This way, there were only four episodes: the decennial in 2011 and revisions in 2015, 2017, and 2019.

The analysis is presented in Table A50. In the table, the columns titled "Full sample" duplicate the values from Table 5 and thus rely on all the episodes that are available. By contrast, the columns titled "Only the decennial" use data from just the decennial episode. By restricting to the decennial episode, we again ensure that the sample includes a single registrant-episode combination per registrant. Thus, by comparing results for the two samples, we can again learn about the influence of stacking.

Table A50 provides further evidence that stacking is not distortionary. The results that rely on stacking ("Full sample") are highly similar to those that do not ("Only the decennial"). Thus, stacking does not appear to be a source of concern.

Table A50: Effects calculated using differences in assigned districts for multiple chambers: with and without stacking

| | **Full sample** | | **Only the decennial** | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Weighted sum of competitiveness in a registrant's districts: all chambers | 1.01*** | | 1.00*** | |
| | (0.185) | | (0.196) | |
| Sum of competitiveness in a registrant's districts: U.S. House | | 2.57*** | | 2.57*** |
| | | (0.774) | | (0.902) |
| Sum of competitiveness in a registrant's districts: state chambers | | 1.12*** | | 1.12*** |
| | | (0.233) | | (0.238) |
| Turnout percentage | 58.0 | 58.0 | 56.0 | 56.0 |
| Clusters | 540 | 540 | 254 | 254 |
| Registrants | 5,604,366 | 5,604,366 | 4,273,537 | 4,273,537 |
| Registrant-episodes | 8,244,415 | 8,244,415 | 4,273,537 | 4,273,537 |
| Registrant-episode-elections | 27,919,274 | 27,919,274 | 21,367,685 | 21,367,685 |

The table shows how the results in Table 5 depend on stacking. Values in the "Full sample" columns match those in Columns 1 and 2 of Table 5 and rely on stacking. Values in the "Only the decennial" columns use just data from the decennial redistricting episode and do not rely on stacking. All other details are the same as in Table 5.

# A11 Assessing linearity

In this appendix, we study an additional feature of the functional form of the relationship between turnout and district competitiveness. We ask whether turnout is linear in the degree of competitiveness or whether the effects of competitiveness vary for higher or lower levels.

Table A51: Linearity in the relationship between turnout and district competitiveness

| | **Main** | **Competitiveness** | | **Polynomial** | |
|---|---|---|---|---|---|
| | | High | Low | Square | Cubic |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 1.38*** | 1.24*** | 1.33*** | 1.48*** |
| | (0.234) | (0.500) | (0.468) | (0.231) | (0.451) |
| Sum of squared competitiveness in a registrant's districts | | | | 0.395 | -0.099 |
| | | | | (1.53) | (2.31) |
| Sum of cubed competitiveness in a registrant's districts | | | | | -3.79 |
| | | | | | (9.96) |
| Turnout percentage | 58.1 | 59.6 | 57.1 | 58.1 | 58.1 |
| F-stat. for joint significance | - | - | - | - | 0.16 |
| p-value for joint significance | - | - | - | - | 0.856 |
| Clusters | 338 | 238 | 209 | 338 | 338 |
| Registrants | 5,203,371 | 3,656,591 | 3,503,669 | 5,203,371 | 5,203,371 |
| Registrant-episode-elections | 31,366,989 | 18,370,399 | 19,492,925 | 31,366,989 | 31,366,989 |

The table displays results from alternative versions of Model (3). "Main" is the main version, as shown in the "All" column of Table 1. "Competitiveness—High" ("Low") restricts the sample to registrants assigned to districts in the top (bottom) three-quarters of the competitiveness distribution. These are districts with competitiveness above (below) 0.657 (0.842). The "Polynomial" columns add variables for the sum of squared or cubic competitiveness. These are instrumented using the sum of squared or cubic assigned competitiveness. In fitting these models, we subtract 0.75 from $c_{i\tau}$ and $c_{a_i}$ before constructing all variables. This way, the coefficient on $C_{i\tau}$ is the marginal effect for registrants in 62.5-37.5 districts, not 100-0 districts. The "F-stat." and "p-value" rows present results from an F-test for joint significance of the sums of squared and cubic competitiveness. Standard errors are clustered by baseline district in the chamber of interest. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We investigate this question by fitting alternative versions of Model (3). The results, presented in Table A51, provide evidence in favor of linearity.

In the columns titled "Competitiveness—High" and "Competitiveness—Low", we fit Model (3) using either registrants assigned to competitive or uncompetitive districts. Competitive (uncompetitive) districts are defined as those in the top (bottom) three-quarters of the competitiveness distribution. If turnout is linear in competitiveness, then effects calculated on the two samples should be the same; if it is not, then they may differ. For instance, if increases in competitiveness matter more among very competitive districts, then effects for the high-competitiveness sample should be larger. The table shows that effects are larger for this sample ($\alpha$ estimate of 1.38 v. 1.24 for the low-competitiveness sample), but the difference is modest and statistically insignificant.

In the columns titled "Polynomial—Square" and "Polynomal–Cubic", we allow turnout to depend on either a square or cubic polynomial in district competitiveness. In the square column, we modify Model (3) by adding a variable equal to the sum of squared competitiveness since redistricting. We instrument for this using an analogous sum of the square of assigned competitiveness. In the cubic column, we further add a variable and instrument for the sum of cubed competitiveness. If the relationship between turnout and competitiveness is non-linear, then the coefficients on the square and cubic terms should be non-zero. The table shows that the estimates for these coefficients are statistically insignificant. In addition, in the "Cubic" column, an F-test for joint significance of the square and cubic terms returns a p-value of 0.86. Thus, turnout appears to be linear in competitiveness.

# A12 Standard errors and randomization inference

In this appendix, we explore robustness in the statistical significance of our results. We first show that statistical significance is similar under alternative ways of clustering standard errors. We then run a permutation test based on randomization inference.

Table A52 displays statistical significance for versions of Model (3) that use alternative clustering strategies. "Main" uses the main strategy, as in the "All" column of Table 1. This is to cluster on a registrant's baseline district in the chamber of interest. The other columns cluster on, respectively: the assigned district; both the baseline chamber-of-interest district and the assigned district; the intersection of these two districts; the intersection of baseline districts for all chambers; the baseline region; and the baseline county. The table shows that standard errors are similar under all cluster definitions (ranging from 0.18 to 0.23) and that they are largest under the main definition. Thus, our choice of the baseline chamber-of-interest district may be conservative.

Table A52: Standard errors for alternative clustering strategies

| | **Main** | **Alternative clusters** | | | | | |
|---|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 1.30*** | 1.30*** | 1.30*** | 1.30*** | 1.30*** | 1.30*** |
| | (0.234) | (0.178) | (0.230) | (0.182) | (0.208) | (0.210) | (0.209) |
| Turnout percentage | 58.1 | 58.1 | 58.1 | 58.1 | 58.1 | 58.1 | 58.1 |
| Clusters | 338 | 373 | 338 | 946 | 514 | 1,285 | 75 |
| Registrants | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 | 5,203,371 |
| Registrant-episode-elections | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 | 31,366,989 |

The table presents results from Model (3) for alternative ways of clustering standard errors. "Main" clusters on a registrant's baseline district in the chamber of interest. It corresponds with the "All" column of Table 1. Column 1 clusters on a registrant's assigned district (i.e., the post-redistricting district in the chamber of interest that contains the registrant's baseline address). Column 2 clusters in two ways based on baseline district in the chamber of interest and assigned district. The number of clusters reported for this column is the minimum over the two groupings. Column 3 clusters on the intersection of baseline district in the chamber of interest and assigned district. Column 4 clusters on the intersection of baseline districts in all chambers. Column 5 clusters on baseline region. Column 6 clusters on baseline county. Other details are as in Table 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In recent years, researchers have highlighted the benefits of running permutation tests based on randomization inference when studying treatment effects (e.g., Athey and Imbens (2017) and Abadie et al. (2020)). These tests allow simulating the distribution of coefficient estimates that would be obtained if the treatment had no impact. The tests operate by calculating coefficient estimates under a large number of random reassignments of the treatment variable. In our setting, permutation tests are particularly useful because they can fully account for the stacking procedure that underlies the empirical strategy.

Table A53: Summary statistics for the coefficient estimates from the permutation test

| **Min** | **Max** | **Percentile** | |
|---|---|---|---|
| | | 2.5 | 97.5 |
| -0.81 | 1.07 | -0.59 | 0.66 |

The table summarizes the distribution of $\alpha$ estimates generated by the permutation test. The distribution contains 200 estimates, corresponding with 200 random reassignments of districts' competitiveness values. See the text for more details on the test.

We implement a permutation test in four steps. First, for each set of districts, we randomly sort the districts' competitiveness values. For instance, for the decennial districts for the U.S. House, District 1 may get the value for District 12, District 2 may get the value for District 6, etc. Second, we merge the re-sorted competitiveness values with the stacked data and calculate new values of

$C_{i\tau}$ and $C_{a_i\tau}$. Third, we run Model (3) and save the estimate for $\alpha$. Fourth, we repeat the process 200 times. The distribution of estimates from this procedure reveals the types of magnitudes we would expect—due to the empirical strategy—if competitiveness had no relationship with turnout.

Results from the test are presented in Table A53. The table shows the maximum and minimum values of the $\alpha$ estimates over the 200 iterations, as well as the middle-95% range. The results suggest that our main estimate has strong statistical significance. Its value of 1.30 is considerably larger than the largest value over the 200 iterations (1.07). In turn, this implies that the p-value of the main estimate is less than 0.005, which is consistent with the p-values based on clustered standard errors.

## A13 Effects on turnout in primary elections

We next explore whether district competitiveness affects turnout in primary elections.

From a theoretical perspective, the relationship between competitiveness and primary turnout is unclear. This is for three reasons. First, both competitive and uncompetitive districts offer incentives for turning out in primaries. In competitive districts, people may want to vote so as to nominate a high-quality candidate who will perform well in the general election. In uncompetitive districts, the primary is the only opportunity to influence who becomes the legislator. Thus, in contrast with general elections, primaries do not feature a strong association between competitiveness and voting incentives. Second, competitiveness may impact primary turnout via spillover effects. Namely, if living in a competitive district causes a general increase in political participation, then, as part of this, it could spur higher turnout in primaries. Finally, primary and general elections may be sufficiently distinct that effects on turnout in general elections do not translate into effects on primary turnout. For instance, Table A54 shows that only 26% of registrants turn out in primaries, versus 48% in midterms and 64% in presidential elections (Table A8). Thus, the type of people whose general-election turnout is impacted by district competitiveness may differ from the type who vote in primaries. Instead, primary voters may be people with a keen interest in politics who turn out regardless of electoral conditions.

Table A54: The effects of district competitiveness on primary turnout

| | **All** | **Chamber** | |
|---|---|---|---|
| | | U.S. House | NC legisl. |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | -0.109 | 0.672 | -0.203 |
| | (0.185) | (0.601) | (0.188) |
| Turnout percentage | 25.6 | 26.5 | 25.3 |
| Clusters | 338 | 36 | 302 |
| Registrants | 5,203,371 | 1,586,751 | 4,653,157 |
| Registrant-episode-elections | 31,366,989 | 6,505,033 | 24,861,956 |

The table presents results analogous to those in Table 1 but where the outcome is turnout in a primary election. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We provide results for effects on primary turnout in Table A54. The results are from a version of Model (3) where the outcome is turnout in a primary, rather than in a general election. The results indicate that district competitiveness has little-to-no impact on turnout in primaries. Effects by chamber have differing signs; the average effect is a precise zero.

Table A55: Effects on general-election turnout for primary voters and non-primary voters

| | **All** | **Voted in baseline primary** | |
|---|---|---|---|
| | | No | Yes |
| Sum of competitiveness in a registrant's districts, $C_{i\tau}$ | 1.30*** | 1.39*** | 0.721*** |
| | (0.234) | (0.264) | (0.216) |
| Turnout percentage | 58.1 | 52.3 | 90.6 |
| Clusters | 338 | 338 | 338 |
| Registrants | 5,203,371 | 4,594,217 | 1,016,090 |
| Registrant-episode-elections | 31,366,989 | 26,497,632 | 4,600,195 |

The table shows heterogeneity in the effect of district competitiveness on general-election turnout by whether a registrant turned out in the primary associated with the registrant's baseline election. Other details are the same as in Table 1.

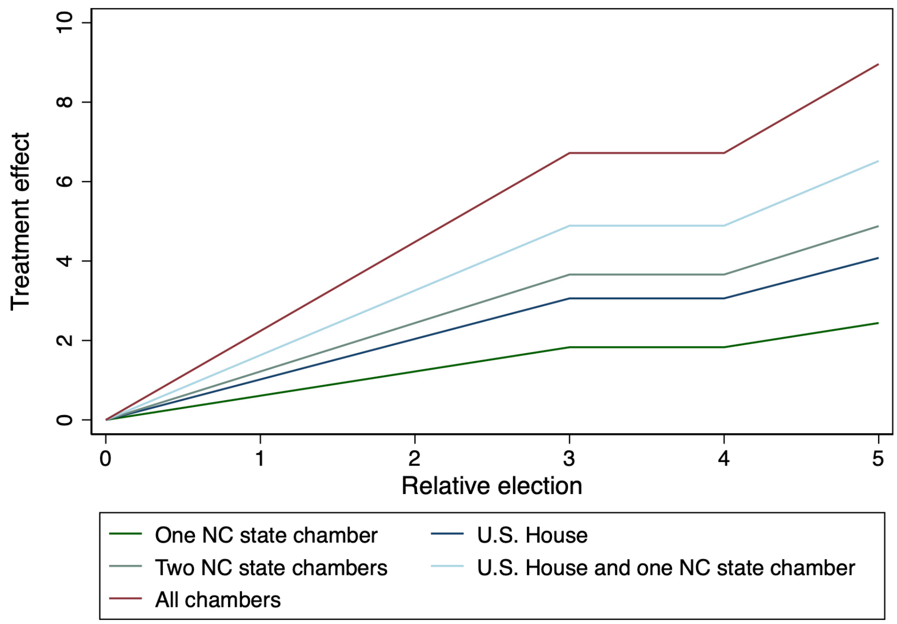
We probe the null result further in Table A55. In the table, we present effects on general-election turnout separately by whether a registrant voted in the primary for the baseline election.

The results allow comparing primary voters with other registrants in terms of turnout rates and sensitivity to district competitiveness. As hypothesized, primary voters are much more likely to turn out in general elections than non-primary voters, with turnout rates of 91% versus 52%. In addition, the effects of district competitiveness are half as large for primary voters (though still significant). An implication is that the people who are impacted by district competitiveness in terms of general-election turnout tend to be people who would rarely vote in primaries absent any effect of competitiveness. Together with the lack of a relationship between competitiveness and primary-voting incentives, this means that district competitiveness may increase general-election turnout without influencing primary turnout. The results are consistent with those of Coppock and Green (2016) that voters often make habits of voting in specific kinds of elections (e.g., general, primary, municipal, etc.).

# A14 Illustration of how effects aggregate

In this appendix, we illustrate how the effects of district competitiveness aggregate. Specifically, we outline the implications of the findings that turnout depends on cumulative competitiveness and that turnout effects are additive across legislative chambers. We show how these findings imply that effects have the potential to become sizable.

Figure A8: An illustration of how the effects of district competitiveness aggregate




The figure displays the time path of treatment effects for the thought experiment discussed in the text. In the experiment, the control group always experiences 80-20 districts. The treatment group experiences 55-45 districts for the first three elections, 80-20 districts for the fourth election, and 55-45 districts for the fifth. "One NC state chamber" is if treatment and control differ in districts for only a single NC state chamber. "U.S. House" is if they differ only for the U.S. House. The other lines are if the groups differ in districts for multiple chambers.

The illustration relates to the following hypothetical situation. We consider a sample of registrants who are randomly placed into either 55-45 districts or 80-20 districts. We call the registrants in the 55-45 districts the "treatment group" and those in the 80-20 districts the "control group". For simplicity, we suppose that there is perfect compliance with treatment assignments; thus, assigned and experienced districts are the same. We imagine that the control group remains in the 80-20 districts for five relative elections following the treatment assignment. By contrast, the treatment group experiences a varying time path of district types. It is in 55-45 districts for the first three elections. Then it is in 80-20 districts for the fourth election. Finally, it returns to 55-45 districts for the fifth.

The results from the hypothetical are depicted in Figure A8. The figure plots treatment effects on turnout for five versions of the treatment. These differ in the legislative chambers for which the treatment and control groups experience differences in districts. "One NC state chamber" is if treatment and control differ in districts for one North Carolina state chamber and share the same districts for the other chambers. "U.S. House" is if the groups differ in districts for the U.S. House and share the same districts for the state chambers. The other versions consider situations where the groups differ in districts for multiple chambers. The treatment effects represent the predicted difference in turnout between treatment and control for a given relative election and version of the treatment. They are calculated in a manner analogous to Equation (4) in the main text. Specifically, for each type of chamber (U.S. House and NC state) and for each relative election (1 to 5), we multiply (a) the difference in cumulative competitiveness between treatment and control for the chamber type in the relative election by (b) our estimate of the causal effect of competitiveness for the chamber type from Table 1. We then sum these products over the two chamber types by election.

The results in the figure indicate that the effects of district competitiveness can be large in certain cases. How large they are depends on the number of elections and legislative chambers in which registrants experience differences in competitiveness. In the first election following treatment assignment, treatment effects are modest, ranging from 0.61 percentage points if treatment and control differ in only one North Carolina state chamber to 2.24 percentage points if the groups differ in all three chambers. By the third election, effects are three times as large, with a range of 1.83 to 6.72 percentage points. If the groups experience no difference in competitiveness in the fourth election, then the treatment effects remain unchanged. By contrast, if the groups once again experience a difference in the fifth election, then the effects resume rising. They climb to between 2.44 and 8.96 percentage points. We cannot comment on how effects change in elections beyond the fifth, as our data limits us to exploring the impacts of lagged competitiveness for only up to four lags (Appendix A8). Nonetheless, the illustration shows that even over five elections—the number for which legislative districts are meant to last—differences in exposure to competitiveness can potentially have a large influence on turnout.

# A15 Comparison with the redistricting literature

There is now a well-developed literature that exploits redistricting to learn about the causal effects of legislative districts. In this appendix, we briefly summarize the existing literature and situate our paper in it.

The literature begins with Ansolabehere, Snyder, and Stewart (2000), who use redistricting to study the role of the personal vote in explaining the incumbent advantage. This paper has county- and town-level data and compares post-redistricting U.S. House vote shares in areas that have the same incumbent as before redistricting with those that have new incumbents. It adjusts for differences in partisanship across areas by controlling for presidential vote shares.

A major advance in the literature was Henderson, Sekhon, and Titiunik (2016). This paper studies the effect of Hispanic incumbents on Hispanic turnout. In doing, it notices that policy-makers strategically place people into districts based on political and demographic factors. As a result, people who are assigned to districts with different characteristics are not comparable.

Henderson, Sekhon, and Titiunik (2016) deal with the lack of comparability by developing a matching procedure. Their data varies at the level of the census block. Thus, their procedure matches entire blocks, not individual people. The procedure has three steps. First, the researchers divide their setting (California) into regions based on the intersection of U.S. House and state legislative districts in the election before redistricting. Second, they impose sample restrictions on regions and blocks. They drop regions where the number of control blocks (those with a non-Hispanic incumbent before and after redistricting) is less than twice the number of treated blocks (those with a non-Hispanic incumbent before and a Hispanic incumbent after). Also, they drop treated blocks with covariates that fall outside the support of same-region control blocks. Third, the researchers create matched pairs of treated and control blocks by matching each treated block to a same-region control block. To do this, they use an inexact matching algorithm that maximizes covariate balance across a variety of block-level variables.

A further advance in the literature was provided by Fraga (2016). This paper develops a matching procedure based on individual-level data, not block-level data. It uses the procedure to explore the turnout effects of a district's racial composition. Moskowitz and Schneer (2019) adapt the Fraga (2016) approach in their study of the short-term effects of district competitiveness. In our paper, we build on the Moskowitz and Schneer (2019) strategy, but we make a number of modifications to facilitate studying long-term effects and effects for multiple chambers. We discuss the modifications in detail in the next section.

# A16 Comparison with Moskowitz and Schneer (2019)

In this appendix, we compare our paper with Moskowitz and Schneer (2019). We first summarize the earlier paper. We then contrast the settings and empirical strategies. Finally, we assess the degree of similarity in the results.

## A16.1 Summarizing Moskowitz and Schneer (2019)

Moskowitz and Schneer study the turnout effects of being assigned to a more competitive U.S. House district as part of the decennial redistricting episode. They use data on registration and turnout during the 2008 to 2014 elections from the data vendor Catalist. Their data includes over 2 million individuals from all 50 states.

Moskowitz and Schneer use a variety of empirical strategies to assess the effect of competitiveness. One of these is a matching strategy that is similar to the one in our paper. When Moskowitz and Schneer use this strategy, they find that being assigned to a more competitive district has a statistically significant effect on the probability of turning out. However, they interpret the magnitude of the effect as "substantively near zero".

## A16.2 Comparing the settings and empirical strategies

The setting of our paper differs from that of Moskowitz and Schneer in important ways. First, Moskowitz and Schneer examine all 50 states, while we focus on North Carolina. Second, Moskowitz and Schneer observe outcomes only in the 2012 and 2014 elections; by contrast, we see them through 2020. Finally, Moskowitz and Schneer study only the U.S. House. In comparison, we consider both the U.S. House and two state legislative chambers.

The empirical strategies used in the papers are broadly similar. Like us, Moskowitz and Schneer exact-match registrants on demographics, pre-redistricting turnout, and pre-redistricting residential location. Also like us, they then assess whether registrants who are assigned to more competitive districts turn out more in post-redistricting elections.

Despite the similarities, there are some technical differences between our approach and theirs. First, in creating match-groups, we match on pre-redistricting districts for both the U.S. House and the state chambers. In addition, we match on assigned post-redistricting districts for chambers other than the chamber of interest. Meanwhile, Moskowitz and Schneer match only on pre-redistricting U.S. House districts. Second, we use a larger number of competitiveness measures. They use only two, one based on the Cook PVI and one based on the ex-post closeness of district races. By contrast, we use three measures of district competitiveness and two measures of race competitiveness. Third, Moskowitz and Schneer scale their measures in a different way. They define the Cook measure as $-100 \cdot |\mathrm{PVI}_d|$. Similarly, they define race closeness as $-100$ times the absolute two-party win margin.[19] Fourth, Moskowitz and Schneer do not instrument for experienced competitiveness using assigned competitiveness. Instead, they restrict the sample to people who stay in their assigned district during the entire analysis period. This way, their first-stage coefficients equal 1 by construction. Finally, Moskowitz and Schneer make a different functional form assumption with respect to the relationship between competitiveness and turnout. They do not model turnout as depending on cumulative competitiveness, $\mathrm{C}_{i\tau}$. Rather, they calculate the average effect of current competitiveness, $\mathrm{c}_{i\tau}$, across both the 2012 and 2014 elections.[20]

19. The absolute two-party *win* margin is half the absolute two-party *vote* margin.

20. When Moskowitz and Schneer instead rely on race closeness, they relate turnout to same-election closeness.

## A16.3 Comparing the results

We now compare our results with those of the earlier paper. In order to understand the source of any differences, we conduct the comparison in three steps.

In the first step, we implement the Moskowitz and Schneer approach in our North Carolina setting, using the same redistricting episode and analysis period as in their paper. Specifically, we use their approach to calculate the effect of competitiveness for U.S. House districts in North Carolina in the 2012 and 2014 elections. We then compare these results with the effects reported in their paper. This first step reveals the impact of the fact that we study just North Carolina, rather than all states.

In the second step, we implement *our* approach on the same sample as in the first step. Specifically, we use our approach to calculate the effect of competitiveness for U.S. House districts in North Carolina in the 2012 and 2014 elections. We then compare these results with the effects based on the approach of Moskowitz and Schneer. This comparison illuminates the impact of differences in the empirical strategy.

In the third step, we implement our approach on our full U.S. House sample. Specifically, we use our approach to calculate the effect of competitiveness for U.S. House districts in North Carolina in all available elections and redistricting episodes. We then compare these results with those computed using just the decennial episode and the 2012 and 2014 elections. This last step shows the influence of using a longer analysis period.

Finally, in all steps, we use the same competitiveness measures as Moskowitz and Schneer. (In some cases, we scale them according to Moskowitz and Schneer's convention; in other cases, we use our scaling.) We do not believe that differences in competitiveness measures are a source of differences in results. This is because, in Table A2, we found that effects vary only slightly across measures.

Table A56: Comparing effects in all states and in North Carolina, using the approach of Moskowitz and Schneer (2019)

| | **All states** | | **North Carolina** | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Cook competitiveness | 0.051*** | | 0.070*** | |
| | (0.012) | | (0.020) | |
| Race closeness | | 0.015 | | 0.025 |
| | | (0.018) | | (0.017) |
| Turnout percentage | - | - | 53.7 | 53.7 |
| Clusters | - | - | 13 | 13 |
| Registrants | - | - | 951,145 | 951,145 |
| Registrant-elections | 2,707,914 | 2,676,278 | 1,902,290 | 1,902,290 |

The table displays results from the first step in the analysis. Specifically, it shows coefficient estimates and standard errors for the coefficient on assigned competitiveness, $c_{a_i}$, in regressions of turnout on $c_{a_i}$ and match-group-by-$\tau$ fixed effects. The results in the "All states" columns are taken from Table 3 in Moskowitz and Schneer (2019). The results in the "North Carolina" columns are obtained by implementing the Moskowitz and Schneer empirical strategy on our data. The sample uses only the decennial redistricting episode for the U.S. House. For outcome data, it uses only the 2012 and 2014 elections. In addition, the sample is limited to registrants who remain registered in their baseline U.S. House district during 2012 and 2014. The model pools observations in the 2012 and 2014 elections and estimates the average effect of $c_{a_i}$ on turnout in these elections. Competitiveness measures are scaled according to the Moskowitz and Schneer convention. "Cook competitiveness" is the Cook PVI-based measure of the competitiveness of the registrant's assigned district. "Race closeness" is the closeness of the district's race in the election in which turnout is measured. Coefficient estimates and standard errors are denominated in percentage points. Standard errors in the "North Carolina" columns are clustered by baseline district in the chamber of interest. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results from the first step are presented in Table A56. They reveal that effects in North Carolina are similar to those calculated using all 50 states. In particular, using a 50-state sample, Moskowitz and Schneer report that a one point increase in their Cook competitiveness measure causes an average increase in turnout in 2012 and 2014 of 0.051 percentage points. In North Carolina, we find that the effect is 0.070 percentage points. For race closeness, Moskowitz and Schneer report a coefficient estimate of 0.015; we find an impact of 0.025.

Table A57: The impact of differences in the empirical strategy

| | **M&S approach** | | **Our approach** | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Cook competitiveness | 2.34*** | | 2.28*** | |
| | (0.680) | | (0.568) | |
| Race closeness | | 0.821 | | 1.18** |
| | | (0.551) | | (0.453) |
| Turnout percentage | 53.7 | 53.7 | 53.7 | 53.7 |
| Clusters | 13 | 13 | 13 | 13 |
| Registrants | 951,145 | 951,145 | 951,145 | 951,145 |
| Registrant-elections | 1,902,290 | 1,902,290 | 1,902,290 | 1,902,290 |

The table displays results from the second step in the analysis. The values in the "M&S approach" columns are isomorphic to those in the "North Carolina" columns of Table A56. However, the coefficient estimates and standard errors are transformed so as to correspond with our convention for scaling the competitiveness measures. As part of this, we multiply the coefficient estimates and standard errors by 2/3. This way, they represent the effect of one election worth of exposure to competitiveness, rather than the average of the effects of one and two elections worth of exposure, as in Table A56. The values in the "Our approach" columns are obtained by implementing our empirical strategy on the same sample as in the "M&S approach" columns. Specifically, they are coefficient estimates and standard errors for the coefficient on $C_{a_i \tau}$ in a regression of turnout on $C_{a_i \tau}$ and match-group-by-$\tau$ fixed effects. Since the sample is restricted to registrants who do not move, this coefficient is equivalent to $\alpha$ in Model (3). See the notes to Table A56 for details on the sample and on clustering of standard errors.

Table A57 provides the results for the second step of the analysis. It suggests that differences in the empirical strategy also have little impact on estimates of causal effects. The first two columns in the table list effects calculated using the Moskowitz and Schneer approach. (The values in these columns are the same as those in the "North Carolina" columns of Table A56. However, they are transformed in accordance with our scaling of competitiveness measures.) The last two columns present effects calculated using our approach on the same 2012-2014 sample. The table shows that the coefficient estimates are similar across the two sets of columns.

Table A58: The impact of differences in the analysis period

| | **2012 & 2014** | | **All elections & episodes** | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Sum of competitiveness in | 2.28*** | | 2.02** | |
| a registrant's districts: Cook measure | (0.568) | | (0.819) | |
| Sum of race closeness in | | 1.18** | | 1.23*** |
| a registrant's districts | | (0.453) | | (0.403) |
| Turnout percentage | 53.7 | 53.7 | 58.3 | 58.3 |
| Clusters | 13 | 13 | 36 | 36 |
| Registrants | 951,145 | 951,145 | 1,586,751 | 1,586,751 |
| Registrant-episode-elections | 1,902,290 | 1,902,290 | 6,505,033 | 6,505,033 |

The table displays results from the third step in the analysis. The values in the "2012 and 2014" columns are the same as those in the "Our approach" columns of Table A57. The values in the "All elections and episodes" columns use all the data that we have available for the U.S. House. The sample for these columns is not restricted to registrants who do not move. Thus, the values in the columns are obtained by implementing Model (3). Standard errors are clustered by baseline district in the chamber of interest.

Table A58 presents results for the third step of the investigation. It indicates that differences in the analysis period also do not affect estimates. In the table, all effects are calculated using our empirical strategy. The values in the first two columns are obtained using just the 2012 and 2014 elections. (They are the same as in the "Our approach" columns in Table A57.) The values in the other columns are computed using all the available elections and redistricting episodes. It can be seen that the coefficient estimates for the full period are similar to those for 2012 and 2014.

The results in this appendix imply that differences between our findings and those of Moskowitz and Schneer are not due to features of our empirical strategy or our setting. Instead, they seem to be due mainly to differences in interpretation. With only two post-redistricting elections, Moskowitz and Schneer didn't realize that the turnout effects of competitiveness grow with additional elections of exposure. In addition, by studying a single legislative chamber, they didn't realize that effects sum across chambers. Consequently, they concluded that competitiveness has little overall impact on turnout. By contrast, we find that effects can add up and thus have the potential to become sizable.

Another implication of the comparison with Moskowitz and Schneer (2019) is that our findings likely have some external validity. Namely, we can replicate Moskowitz and Schneer's results, even though our data is only for North Carolina. This suggests that effects in North Carolina may be representative of those in the country as a whole. A caveat is that the replication is just for 2012 and 2014; as such, we cannot prove generalizability in later elections. This caveat is important because the later part of the sample period is when North Carolinians experienced unusually frequent changes in their districts.

# References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2020. "Sampling-Based versus Design-Based Uncertainty in Regression Analysis." *Econometrica* 88 (1): 265–296.

Ansolabehere, Stephen, James M. Snyder, and Charles Stewart. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage." *American Journal of Political Science* 44 (1): 17–34.

Athey, Susan, and Guido W. Imbens. 2017. "Chapter 3 - The Econometrics of Randomized Experiments." In *Handbook of Field Experiments,* edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:73–140. Handbook of Economic Field Experiments. North-Holland.

Coppock, Alexander, and Donald P. Green. 2016. "Is Voting Habit Forming? New Evidence from Experiments and Regression Discontinuities." *American Journal of Political Science* 60 (4): 1044–1062.

Fraga, Bernard L. 2016. "Redistricting and the Causal Impact of Race on Voter Turnout." *The Journal of Politics* 78 (1): 19–34.

Fraga, Bernard L., Daniel J. Moskowitz, and Benjamin Schneer. 2021. "Partisan Alignment Increases Voter Turnout: Evidence from Redistricting." *Political Behavior.*

Henderson, John A., Jasjeet S. Sekhon, and Rocio Titiunik. 2016. "Cause or Effect? Turnout in Hispanic Majority-Minority Districts." *Political Analysis* 24 (3): 404–412.

Moskowitz, Daniel J., and Benjamin Schneer. 2019. "Reevaluating Competition and Turnout in U.S. House Elections." *Quarterly Journal of Political Science* 14 (2): 191–223.