# APPENDICES FOR ONLINE PUBLICATION

# APPENDIX A: DATA

Ideally, we would check for firm gerrymandering in every state for every districting plan produced. However, data availability limits the set of plans we can analyze. Our simulation procedure requires the following data:

1. Precinct shapefiles
2. Precinct-level vote outcomes
3. Precinct-level population and minority population

Ideally, the precinct-level vote outcome data would include the vote share for the House of Representatives or state legislators since those are the elected officials most impacted by redistricting. However, that data is rare. For most states, precinct-level election data includes presidential election results and possibly senate or gubernatorial results. For 2012, only Wisconsin, Maryland, and Minnesota make available House of Representatives data that are linked to precinct shapefiles. As well, even in states where House of Representatives data are available, many House elections are often uncontested, making those results useless for estimating latent two-party vote share.

To overcome this difficulty, we calculate an average two-party vote share using the closest available elections. We detail which elections we use for each state below.

This procedure (averaging) helps mitigate concerns about the district invariance assumption. Furthermore, in Appendix D, we will show that our results are robust to swings in vote shares, which effectively tests the district invariance assumption.

Precinct-level data availability is much greater for more recent elections; nearly all states have precinct-level data available for 2020. We plan to replicate our results for all 43 states that redistrict once 2022 precinct-level results are available, which will allow greater leverage to study how commissions or changes in party control affect firm allocations.

**Table A1: Elections aggregated for vote outcomes.**



# APPENDIX B: SIMULATIONS

Our simulations leverage the Sequential Monte Carlo (SMC) approach of McCartan and Imai (2020), which “generates a sample of redistricting plans converging to a realistic target distribution,” the null distribution of legislative districting plans. The advantages of this method are threefold. First, it enables efficient, parallel sampling of a complex distribution space, allowing us to create tens of thousands of potential maps in a short amount of time. Secondly, SMC is the only simulation method to allow flexible sampling constraints that can fit real-world legal criteria for redistricting, including VRA compliance, compactness, maximum count of county or city splits, and maintaining the cores of existing districts. Third, SMC is the only approach that is validated both empirically and legally: through a series of cases where researchers can enumerate the entire set of possible districts, SMC samples representatively from that space while preserving legal criteria; as well, in litigation SMC has generated substantial evidentiary support.

We note in our Methods section that our procedure follows five steps:

1. Generate a precinct adjacency matrix from a state precinct map.
2. Conduct the Sequential Monte Carlo sampler, drawing 50,000 plans per state.
3. Overlay each generated plan with our data set of geolocated firms.
4. Infer expected two-party vote shares for every district in every plan.
5. Finally, we assign firms to parties based on district vote shares.

Here we detail our methodological choices in Steps 1 through 5.

A2.1 Generating an Adjacency Matrix

The first step in our simulation procedure is to generate a precinct adjacency matrix for every state in our sample. For most states, we perform this step using the “redist\_map” function from the redist library in R, verifying using the “check.contiguity” function from the geomander library to ensure that our adjacency matrices are complete.

However, for a number of state shapefiles (Maryland, Ohio, Wisconsin) we perform additional manual preprocessing to correct imperfect shapefiles. For these cases, we manually add edges between precincts using the “add.edge” function. In cases of an island precinct with no connections to any mainland precinct, we select the three closest mainland precincts and add edges to those.

A2.2 Running the SMC Sampler

We run the SMC procedure using the “redist\_smc” function from the redist library. The key inputs to this function are the adjacency matrix from A3.1, a vector indicating the population of each precinct, the number of districts to allocate precincts into, the number of simulated plans to create, and the population tolerance: the maximum allowable deviation from equal population in each district. In reality this population deviation must be strictly minimized; we use a tolerance of 1% for computational efficiency. We also compute 50,000 plans for each state-year.

Like all simulations, SMC involves a number of hyperparameters which affect the speed, efficiency, and validity of the simulated districts. One such hyperparameter is a vector identifying which county each precinct falls into; if provided, the SMC function tries to minimize county splits, but providing it also helps to speed up the sampler. We provide this vector whenever it is available.

The remaining hyperparameters relate to simulation constraints, and in all cases, we follow best practices as outlined by the authors of the redist library. We add a constraint for compactness, set to a strength of 1, such that more compact districts are preferred. For states that are bound by the Voting Rights Act, we also add a set of constraints to accommodate the requirement that majority-minority districts be created if possible. For this constraint we supply a vector enumerating the minority population in each precinct, a threshold of 0.55 indicating that a majority-minority district must be created if it is possible to reasonably produce a district with 55% minority population, and a strength parameter, which we vary across states: we increase the strength of this parameter until it begins to reduce other performance metrics of the sampler.

We design these constraints to be conservative in the creation of majority-minority districts. The redist authors suggest that to set the strength parameter in the VRA constraint, test increasing the value until the sampling efficiency drops to less than 75%. We follow this guidance, resulting in strength parameters that range from 5 to 100. Notable is that the redist authors often use strength parameters of 1,000 or more.

A2.3 Overlaying Maps with Firms

The output of the previous step is a set of 50,000 precinct assignments: which district each precinct falls into for each hypothetical map. This makes it easy to assign firms to districts: since firms fall neatly into precincts, it is possible to merge the firm location data set with the precinct assignment data set without requiring any spatial merge functions.

A2.4 Inferring Two-Party Vote Share

To determine which party is likely to control each hypothetical district in each hypothetical map, we calculate a measure meant to capture overall latent Democratic vote share for each precinct. Ideally, we would use the House of Representatives election immediately following the introduction of new maps, but that precinct-level data is not always available, and in some cases, races may be uncontested. Therefore, we average both the Democratic and Republican vote shares for a number of available elections to produce a more robust measure of latent Democratic preference at the precinct level. For example, in Oregon we average the Democratic and Republican vote counts for the President, House, Governor, Attorney General, and Secretary of State in 2016. For the complete details of which elections we aggregate for each state, see Appendix A.

Having produced these expected vote counts, we aggregate our data set to the district-map level by summing the expected Democratic and Republican votes within each hypothetical district, then determining which party has more latent support in those hypothetical districts.

Assigning Firms to Parties

Finally, we are left with a data set of 50,000 sets of districts, the count of how many firms are in those districts, and the latent Democratic vote share for each district. We can then aggregate up to the map-level and calculate how many firms are won by Democrats in each simulated redistricting plan.

**Plan Representativeness**

Our goal is to use this set of 50,000 plans as a null distribution that is *representative* of the set of legally-valid redistricting plans. Earlier tools for computer-derived redistricting plans could not meet the bar of representativeness: “redistricting remains an extremely difficult computational problem. No algorithm is known to exist that produces optimal plans for any redistricting problem of realistic size, and because of the sheer mathematical complexity of redistricting, it is unlikely that the computational problem of redistricting will be solved, at least for tasks such as redistricting large states at the block level” (Altman et al. 2005). More recent debates have argued that such methods are not “sampling from a specified target distribution”, and that they can only conduct “local explorations of the space of redistricting plans, and could not therefore have generated a representative sample of all valid plans” (McCartan & Imai 2023).

However, the SMC method has made excellent progress since 2005 and can “obtain a representative sample of alternative redistricting plans under [given] redistricting criteria” (McCartan et al. 2022), and implement “several diagnostics” in order “to ensure that the sampler is not getting stuck, or producing an unrepresentative sample.” The SMC method specifies a valid target distribution, and includes built-in diagnostics for assessing sample diversity. Most importantly, McCartan and Imai (2023) perform a validation study using a case where all possible district plans are enumerable, and show that SMC obtains a representative sample of all plans across a variety of statistical and substantive estimands.

In short, the most recent evidence strongly indicates that unlike previous methods, SMC can “incorporate each state’s specific requirements for map drawing, along with federal requirements, to ensure that the sample of simulated plans is representative of the space of legal plans” (Kenny et al. 2023).

**State Legislative Districts**

We cannot yet study firm gerrymandering at the state legislative level. We have precinct results for statewide contests and for Congressional Districts, but not for state legislative districts. We could use Congressional or statewide results as a proxy for state legislative vote, but that would entail making some very strong assumptions. We know from Kuriwaki’s (2020) analysis of raw ballot images that voters split their ticket between the governor and their state legislative vote as much as 30% of the time. This is high enough inconsistency that we are uncomfortable substituting gubernatorial vote for state legislative vote.

# APPENDIX C: Robustness

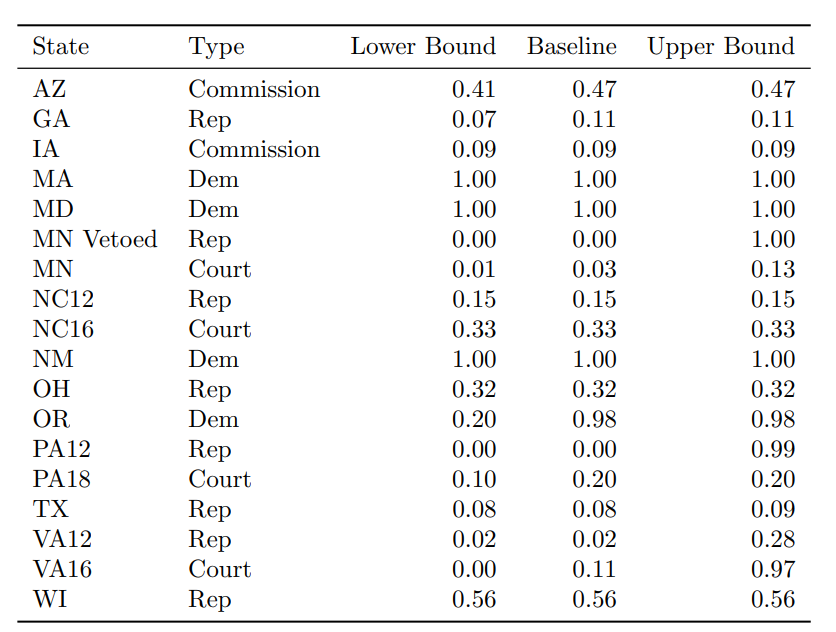
## Maps that are robust to swings in vote share

Our analyses thus far assume that mapmakers draw lines with foreknowledge of how a district will vote. For most cases this assumption is safe: mapmakers have detailed data on the districts they are drawing, and only 31 Congressional elections, or 7.1%, were won by a margin of less than 5 percentage points in 2012. But if mapmakers lose a district that they have drawn expecting to win or vice versa, then their gerrymanders could appear either less effective or more effective than they intended. More generally, mapmakers want to both win as many seats as possible and to win them comfortably, two goals that often conflict. Drawing many districts that they expect to win narrowly may backfire during a swing in voter preferences (Gul & Pesendorfer 2010).

To capture this consideration, we conduct a bounding exercise. We imagine 5 percentage-point uniform shocks to vote share in both directions and reexamine firm gerrymandering in these extreme scenarios: at the lower bound, the outparty wins all districts won by a margin slimmer than 55-45; at the upper bound, the mapmaker’s party wins them.[[1]](#footnote-1) If firms appear to be gerrymandered in favor of the mapmaker’s party even when the outparty gains an extra 5% of the vote, we can conclude that firm gerrymandering is not the result of swing districts and secular shocks to vote share.[[2]](#footnote-2)

Table 2 displays quantile values for these bounds relative to each state’s counterfactual distribution. Column 2 contains reference values for firm quantiles obtained in our core analysis, presented visually in Figure 3. Column 1 displays the lower bound from the exercise, and Column 3, the upper bound.

**Table A2 - Quantiles of Firms Allocations by State from Bounding Exercise**

  
Notes: The lower bound is the quantile of the enacted plan when Republicans are given all of the firms located in marginal (45-55) districts. The baseline is the quantile of the enacted plan when the party with the simple majority is given firms in marginal (45-55) districts, as in Figure 3. The upper bound is the quantile of the enacted plan when Democrats are given all of the firms located in marginal (45-55) districts.

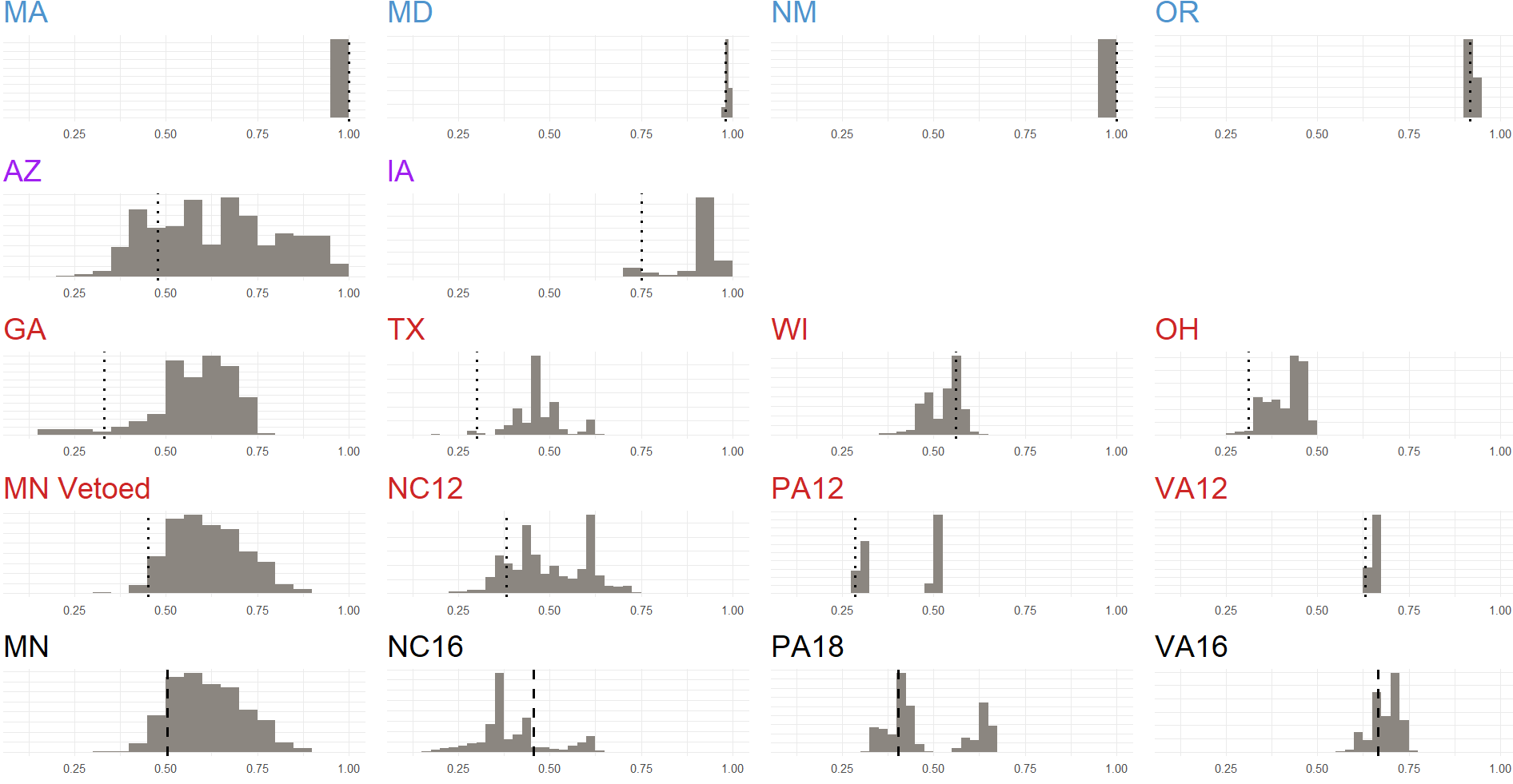
This analysis exposes how robust the mapmakers’ choice of a redistricting plan is relative to a counterfactual world in which firms in all other potential redistricting plans are allocated by simple majority. In the majority of states (e.g. Iowa, Maryland, Massachusetts, New Mexico, North Carolina, Ohio, Texas, and Wisconsin) both bounds are equivalent to the reference point because mapmakers enacted a plan with no firms in marginal seats. In other states, one bound may equal the baseline, while the other may not, reflecting that firms only reside in marginal districts of one party in the enacted plan. In a few states (e.g. Minnesota, Oregon, Pennsylvania, and Virginia), we see a large range in quantiles obtained at different bounds, indicating that a large number of firms reside in the few districts we classified as marginal. The quantity of firms in marginal districts could represent geographic constraints or mapmakers’ appetite for risk.[[3]](#footnote-3)

Taken together, these results show that our finding that firms are gerrymandered does not hinge on assumptions we made about the treatment of marginal seats. Following Caughey et al. (2017), as above, we calculate a global p-value for the bounding analysis of 9.55 x 10-5, indicating strong statistical evidence that firm gerrymandering exists in at least some states, even when actual election results bias against the existence of firm gerrymandering in enacted plans.[[4]](#footnote-4)

**Conditional Gerrymandering results**

Figure A2 presents figures equivalent to Figure 2, our grid of firm gerrymandering histograms, but examining conditional gerrymandering instead. As the case studies demonstrate, the statistical tests in the conditional distribution are underpowered in some cases, as the range of firms in any given seat distribution can be small and discontinuous. That said, many of the states that appear as strong unconditional gerrymanders also appear as strong conditional gerrymanders (e.g., Virginia, Georgia, and Pennsylvania, for Republicans, and to a lesser extent, Massachusetts and New Mexico for Democrats). Several states which do not appear as strong unconditional gerrymanders appear as statistically significant conditional gerrymanders favoring Republicans: Texas, Ohio and Iowa are in this category. Ohio, in particular, moves from being quite proportional in the unconditional distribution to a fairly strong Republican firm gerrymander in the conditional distribution. The somewhat partisan Iowa commission, which is subject to a legislative veto, disproportionally allocated firms to Republicans in the conditional distribution, consistent with the partisan composition of the legislature. North Carolina is nearly an unconditional firm gerrymander favoring Republicans, but it is not a firm gerrymander conditional on seats, which tilt Republican. Wisconsin is the only partisan state that is not close to being a gerrymander in the conditional or unconditional distribution. Wisconsin Republicans seem to have ignored firms relative to other criteria when drawing congressional district lines. On the Democratic side, Maryland is similar to Oregon: It is clearly an unconditional firm gerrymander favoring Democrats, but because an extreme seat distribution overlaps with an extreme firm distribution and the sample of firms in the conditional distribution is small, it is impossible to separate seat and firm gerrymandering.

**Figure A2: Plans drawn by Republicans (Democrats) put fewer (more) firms in Democratic districts compared to simulated plans, using the conditional distribution. Court-drawn plans (bottom row) are closer to the conditional mode than the original enacted plans in those states.**



APPENDIX D

Our main results show that the mapmaker’s party obtains an unexpectedly high share of firms in general, and that when we condition on the number of seats, we can sometimes show that mapmakers target firms in particular. One might wonder, however, whether there are other confounding variables besides seat share that are driving our results. In particular, if there are things that mapmakers desire that are spatially collocated with firms, our current analysis might not disentangle firms from those confounders.

We address this concern in two ways. First, we show that other precinct features mapmakers might specifically target are not particularly collocated with firms. Second, we conduct a cross-state regression analysis of unusual precincts (i.e., those that appear to be gerrymandered specifically), confirming that precincts containing firms are targeted in some states even when controlling for some other potential objects of desire for politicians: neighborhoods with particular racial demographics, urban neighborhoods, and wealthy residents.

Our dependent variable in these regressions is a measure of “unusual” precincts, which we define as those precincts that belong to the “wrong” party. Specifically, these precincts belong to the Democratic (Republican) party in 70% or more of the simulations, but belong to the Republican (Democratic) party in the enacted plan. Just under 10% of precincts count as “unusual” under this threshold; we view these as precincts that are most likely to have been politically manipulated.

Our key independent variable is a firm dummy that takes value 1 if there is a publicly traded firm in that precinct. By adding other precinct-level variables such as the count of wealthy voters, the urban population, and the racial demographics as controls, we can show that the presence of firms predicts whether a precinct belongs to the “wrong” party, independent of these other features that mapmakers might target.

**Co-location of Firms and Other Objects of Desire**

First, we present a descriptive analysis of unusual precincts by state in Table A3 and Table A4. These tables show four important things.

1) Among unusual precincts, whether a precinct has firms is essentially uncorrelated with key demographic variables about that precinct (no higher than 0.2 for income, and often negative for urban and white, shown in Table A3).

2) At the descriptive level, consistent with our outlier analysis, firms located in unusual precincts overwhelmingly go to the mapmaker’s party. To wit: all 9 firms in unusual precincts in Democratic states go to Democratic districts; while 202 of the 268 firms located in unusual precincts in Republican states go to Republican districts.

3) A regression analysis is not feasible for most states. Massachusetts, New Mexico, Arizona, Virginia and Wisconsin have no firms in unusual precincts. In Democratic Maryland and Oregon there are only 3 and 6 firms in unusual precincts, all won by Democrats, while in Republican Pennsylvania there are 42 firms in unusual precincts, all won by Republicans.

4) In several of the remaining large Republican states of our sample (Texas, Georgia, North Carolina and Ohio), descriptive statistics of unusual precincts also point to firms being disproportionally allocated to the mapmaker’s party (e.g., in Georgia the count is 22-6 in favor of Republicans, while it is 16-6 in NC, 17-12 in Ohio and 121-48 in Texas). These states provide enough variation to make reasonably valid inferences from a regression analysis.

5) The cleanest commission state—Arizona—has no firms in any of its 127 unusual precincts.

**Table A3. Correlation matrix, firms and other variables in unusual precincts in**

**states with >6 firms located in unusual precincts**

|  |  |  |  |
| --- | --- | --- | --- |
|  | High Income | Urban | White |
| GA | 0.2 | -0.08 | 0.04 |
| MN | 0.09 | -0.11 | -0.14 |
| NC | 0.17 | -0.1 | 0.05 |
| OH | 0.04 | 0.23 | -0.06 |
| PA | 0.1 | -0.03 | -0.08 |
| TX | 0.19 | -0.1 | 0.13 |

**Table A4. Precinct Level Descriptive Statistics**



**Unusual Precinct Regressions**

Next, we present the results from a series of regressions designed to show robustness to a number of important confounders. The dependent variable is our measure of unusual precincts based on a measure of the likelihood that a precinct is placed differently in the realized plan than in our 50,000 simulations in each state. As mentioned above, it is a categorical variable with three values. It takes the value of 1 if the precinct usually goes to the Democratic party in our simulations but is allocated to a district won by Republicans in the realized plan. It takes a value of 2 if the precinct goes to the Republican party in our simulations but is allocated to Democrats in the realized plan. It takes 0 otherwise, meaning that the precinct in the enacted plan is not unusual: it belongs to the same party to which it is usually allocated in our simulations.

In the benchmark regression we use a cutoff of 70%: a precinct is considered unusual if it is allocated more than 70% of the times in our simulations to a certain party but goes to the other in the realized plan. At this threshold, 10% of the precincts are classified as being unusual. Our results do not depend on the specific threshold, as we obtain similar conclusions when using different reasonable thresholds such as one-third, one-fourth, or one-tenth. Overall, this variable is a proxy for the extent of manipulation by mapmakers. Unusual precincts are the ones more likely to have been gerrymandered deliberately. Hence, by construction, we are only focusing on those precincts that are allocated to a party that is different from what would be expected *given the voting behavior of the precinct (which is determined by voter characteristics)*. This means that a precinct would not be considered as unusual if, for example, its voting history and the geographical location made it equally likely to be allocated to Republicans or Democrats by our simulations or if it were to allocated to a Republican or Democratic district “usually” by our simulations. Only precincts that are allocated in an unusual fashion given their voting behavior and location are considered as suspects of gerrymandering in our empirical analysis. We believe that this empirical strategy sets up a conservative bar for detecting whether firms are an object of desire, since mapmakers might also have been interested in firms that are located in “not-unusual” precincts, but did not need to go the extra mile to grab those.

The right-hand side variables are the count of firms (our key predictor) as well as the racial composition, urban status, and number of high-income individuals in the precinct, which are all available alternative factors that might explain which precincts are targeted, plus state fixed effects.[[5]](#footnote-5) The racial composition of the precinct is obtained from the Census, and is measured at the precinct level. The urban status of the precinct is defined using the urbanization variable from the Census, which is a variable that takes values from 0 to 8, with 8 being precincts in the most urban areas and 0 in the most rural areas. Income is defined as the number of voters in the top income bracket (+200,000$) in the precinct. The Census variables, however, are measured at the ZIP code level rather than at the precinct level. In order to create a measure at the precinct level, we calculate a precinct level weighted average of the ZIP code level variable: when a precinct crossed several ZIP codes, we calculated the precinct level value as the weighted average of the ZIP codes crossed by the precinct, using as weights the percentage of geography of the precinct that corresponded to each of the ZIP codes. This might result in some measurement error, which we assume is random across firm locations.

As explained above, we run the regression analysis only on Republican states where there are firms in each category of the unusual precincts variable (TX, GA, NC and OH). Combined these states have more than 25,000 precincts, and considerable racial, partisan and firm diversity, particularly TX.

We run a series of multinomial logit regressions in which we estimate the probability of a precinct being unusual (unusually Republican, unusually Democrat, or not unusual, which is the benchmark). If firm targeting exists, even conditional on precinct characteristics, we would expect the firm count variable to be positive and statistically significant among partisan states and null among commission- or court-drawn states.

We expect to find that the presence of firms in a precinct positively predicts mapmakers’ interest in allocating them to co-partisan districts, regardless of how likely it is that precinct would naturally fall into a co-partisan district in our simulations. We also might expect to find significant associations for our control variables, suggesting other avenues for research into gerrymandering.

As a placebo outcome, we also analyze whether firms predict precincts that unusually fall into Democratic districts. As our sample consists only of Republican states, we do not expect firms to be predictive here: these unusual precincts are ones that Republican mapmakers appear to have given to Democratic districts, and so these likely contain things that are either (a) toxic to both Republicans and Democrats, or (b) toxic to Republicans but not Democrats.

Column 1 in Table A5 confirms our expectations.[[6]](#footnote-6) While the coefficient that captures wealthy voters is significant and has the expected positive sign in column 1, the relevant coefficient for precincts that contain firms is also positive and significant. Likewise, the coefficient in our placebo regression is very close to zero and not statistically significant. We take these results as evidence consistent with the conclusion that conditional firm gerrymandering exists, at least in some states, and is unlikely to be systematically explained away by potential confounders.

**Table A5 – Precinct-level Regression of Unusual Precincts**



*Note: The dependent variable in this regression takes value 1* *when in 70% or more of simulations the precinct is allocated to Democrats but in the realized plan the precinct is allocated to a district won by Republicans*. *The variable takes value 2 when in 70% or more of simulations the precinct is allocated to Republicans but in the realized plan the firm is allocated to a district won by Democrats. The benchmark is 0, which represent cases in which a precinct is allocated in the realized plan where our simulations usually allocate it. Wald-tests show that Republican and Democrats are generally distinct from each other.*

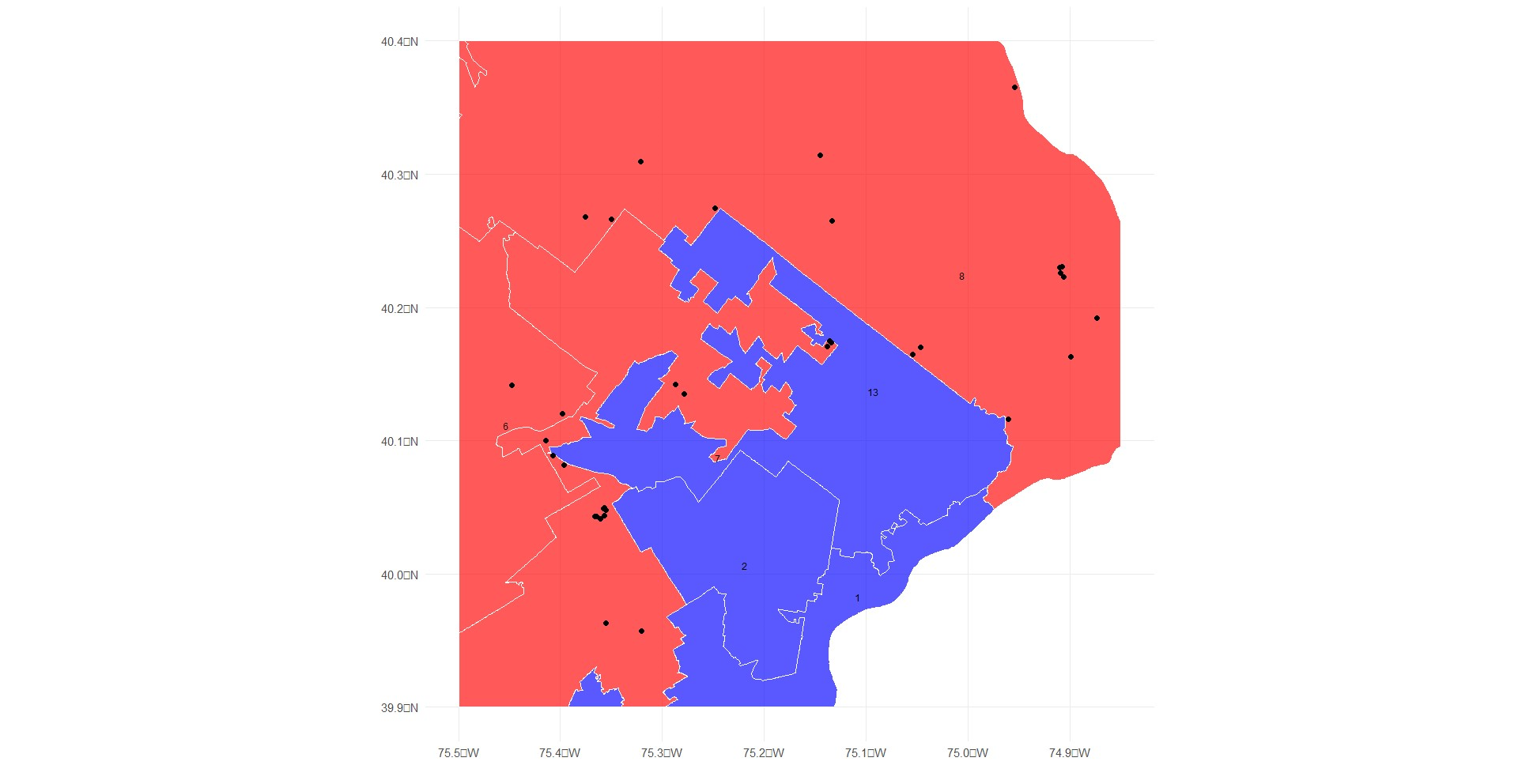
**APPENDIX E**

This appendix presents descriptive statistics of the characteristics of firms in our sample and of the precincts in which they are located, both overall and by state. Analogously to what Appendix 4 showed at the precinct level, we also present a regression analysis with the same states as earlier (Texas, Georgia, Ohio, and North Carolina) in which we try to identify whether certain types of firms are more likely to be gerrymandered than others.

Our sample of firms is obtained from COMPUSTAT, which reports financial information for all firms that are publicly traded in U.S. stock markets. Publicly traded firms are relatively large in terms of sales and employees, which makes them likely to be powerful economic players at the state level. This is an advantage, because there is a wealth of financial data for these firms, but also a disadvantage, because there might not be enough heterogeneity in terms of size or other characteristics: all firms in COMPUSTAT might be above the size threshold that makes them visible and desirable for politicians.

The basic characteristics of our sample are summarized in Table A6. We show results separately for the whole sample of firms in each state and by “unusual” precincts, defined as we did in Appendix 4: a precinct is considered unusual if it is allocated more than 70% of the times in our simulations to a certain party but goes to the other in the realized plan. The descriptive statistics do not show any clear pattern of firms’ traits being more preferred than others. There are many states in which the median profitability or size of firms unusually assigned to the mapmaker’s party are smaller than either that of firms not unusually allocated or than those unusually allocated to the other party. A plausible interpretation is that what determines which firms are grabbed by mapmakers is simply the opportunity to do so. While there might be exceptions, all publicly traded firms seem to be valuable. This opportunism is suggested by the simple visual inspection of unusual precincts in some cases, such as Pennsylvania, as Figure A3 illustrates. This Figure shows the location of firms located in unusual precincts. One can see that many of them are very close to the boundaries of the out-party congressional districts, but none of them actually fall in the out-party congressional district. Specifically, four firms are within 200 meters of a Democratic Congressional district; and an additional 5 firms are within 500 meters of one.

**Figure A3. Location of firms in unusual precincts in the Philadelphia area, with Congressional District boundaries**



Using the variables shown in Table A6, and precinct characteristics, we then perform a regression analysis aimed at detecting systematic differences between the characteristics of firms located in unusual precincts conditional on precinct racial and income demographics. [[7]](#footnote-7) The firm characteristics included in the regression are size (sales in Millions of USD), profitability (earnings over total assets) and innovativeness (proxied by a dummy that takes value 1 if the firm reports positive R&D spending and 0 otherwise). As a first cut, we use these variables because out the variables readily available on COMPUSTAT, they are ones that should matter if politicians care not only about having firms in their districts but about having certain firms with good traits.

The regression results in Table A7 confirm that none of these important firm-level characteristics explain the allocation of firms to gerrymandered precincts. We do not find any clear evidence of any firm characteristics like size or profitability explaining the probability of the firm being gerrymandered into a mapmaker’s district or away into the other party’s district. We note, however, that our firm level analysis should be interpreted with caution, as it is limited to four states (Texas, Georgia, Ohio, and North Carolina), by the number of firms, and by the available set of firms’ characteristics. We expect a much larger sample and more variables would reveal some heterogeneity in the relative value of firms and in terms of the firm-party and perhaps even industry-party match.

**Table A6. Firm Level Descriptive Statistics (**averages for all firms in each state**)**



***Note: The table shows the average value in each cell.***

**Table A7 – Firm-level Multinomial Logit Regression on Attributes Selected**



*Note: The dependent variable in this regression takes value 1 when in 70% or more of simulations the precinct is allocated to Democrats but in the realized plan the precinct is allocated to a district won by Republicans. The variable takes value 2 when in 70% or more of simulations the precinct is allocated to Republicans but in the realized plan the firm is allocated to a district won by Democrats. The benchmark is 0, which represent cases in which a precinct is allocated in the realized plan where our simulations usually allocate it.*

|  |  |
| --- | --- |
|  | |
| ADDITIONAL REFERENCES Altman, M., MacDonald, K. and McDonald, M., 2005. From crayons to computers: The evolution of computer use in redistricting. Social Science Computer Review, 23(3), pp.334-346.  Kuriwaki, Shiro. 2019. Party Loyalty on the Long Ballot: Is Ticket Splitting More Prevalent in State and Local Elections?  McCartan, Cory, Chris Kenny, Tyler Simko, T Garcia III, G., Wang, K., Wu, M., Kuriwaki, S. and Imai, K., 2022. Simulated redistricting plans for the analysis and evaluation of redistricting in the United States. Scientific Data, 9(1), p.689.  McCartan, C. and Imai, K., 2020. Sequential Monte Carlo for sampling balanced and compact redistricting plans. arXiv preprint arXiv:2008.06131. | |
|  |

1. In this analysis, we leave the counterfactual distribution unchanged, i.e. it represents firm allocations going to the party with a simple majority in each district of each redistricting plan. Leaving this counterfactual unperturbed will lead to a relatively conservative test. [↑](#footnote-ref-1)
2. We define marginal seats as those with a predicted district vote share falling within the 45% - 55% interval. Defining marginal seats as those in an interval this wide will lead to relatively conservative results in our bounding exercise. [↑](#footnote-ref-2)
3. We note that the choice of an enacted plan is endogenous to the mapmakers’ appetite for risk: risk-tolerant mapmakers may focus on the lower bound, while risk averse mapmakers may focus on the upper bound subject to other tradeoffs. [↑](#footnote-ref-3)
4. In constructing this global p-value, we assume: when Democrats draw maps, they lose all marginal districts to Republicans; when Republicans draw maps, they lose all marginal districts to Democrats; and, when commissions draw maps the election outcome falls farthest from neutrality given marginal district outcomes. [↑](#footnote-ref-4)
5. We do not include vote share, as the left-hand side variable is definitionally related to expected vote-shares. [↑](#footnote-ref-5)
6. We show the pooled regression with state-fixed effects. Separate regressions for each state shows an independent and significant association in TX and GA, an association consistent with the hypothesis slightly below conventional statistical levels for OH and no association for NC, results that are very consistent with the simulations. [↑](#footnote-ref-6)
7. We could not include urbanization as a control because the models did not converge. [↑](#footnote-ref-7)