Redistricting Reforms Reduce Gerrymandering by Constraining Partisan Actors

Cory McCartan
Department of Statistics
Pennsylvania State University

Christopher T. Kenny
Department of Government
Harvard University

Tyler Simko
Department of Government
Harvard University

Emma Ebowe
Department of Government
Harvard University

Michael Y. Zhao
Harvard College

Kosuke Imai*
Department of Government
Department of Statistics
Harvard University

July 10, 2024

Abstract

Political actors frequently manipulate redistricting plans to gain electoral advantages, a process commonly known as gerrymandering. To address this problem, several states have implemented institutional reforms including the establishment of map-drawing commissions. It is difficult to assess the impact of such reforms because each state structures bundles of complex rules in different ways. We propose to model redistricting processes as a sequential game. The equilibrium solution to the game summarizes multi-step institutional interactions as a single dimensional score. This score measures the leeway political actors have over the partisan lean of the final plan. Using a differences-in-differences design, we demonstrate that reforms reduce partisan bias and increase competitiveness when they constrain partisan actors. We perform a counterfactual policy analysis to estimate the partisan effects of enacting recent institutional reforms nationwide. We find that instituting redistricting commissions generally reduces the current Republican advantage, but Michigan-style reform would yield a much greater pro-Democratic effect than types of redistricting commissions adopted in Ohio and New York.

Keywords redistricting • differences-in-differences • formal modeling • continuous treatment

1 Introduction

Democratic institutions play a crucial role in preventing political actors from pursuing policies that prioritize their own interests over the broader public good (Federalist papers 53 and 54: Madison, 1788b,a). As political polarization increases and democratic norms erode, scholars have recognized the growing importance of institutional constraints on policy makers (Levitsky and Ziblatt, 2023; Little and Meng, 2023; McCarty, 2019). To safeguard institutions, advocates have suggested reforms which insulate democratic processes from partisan control.[1] For example, the Electoral Count Reform and Presidential Transition Improvement Act of 2022 represents an effort to make it more difficult for partisan actors to manipulate the presidential electoral certification process.

*To whom correspondence should be addressed. Email: imai@harvard.edu. Website: https://imai.fas.harvard.edu/. Address: 1737 Cambridge Street, Cambridge, MA 02138. Acknowledgements here.

[1]For example, see People Not Politicians Oregon or Ohio's Citizens Not Politicians.

This paper studies American legislative redistricting, a political process often exploited by partisan actors to enact districting plans that favor their own party. This manipulation, known as *partisan gerrymandering*, has been widespread in the past two redistricting cycles (Arora et al., 2024; Kenny et al., 2023; Warshaw et al., 2022). Redistricting plans that disproportionately favor a certain party can limit how responsive a party's share of seats in the legislature are to changes in its vote share, and can reduce the electoral power of racial minorities (e.g. Canon, 2022; Canon and Race, 1999; Grofman and Handley, 1991; Polsby and Popper, 1991).

Reform efforts to limit gerrymandering are often designed to constrain partisan map drawers. They include the establishment of independent map-drawing commissions and the introduction of court oversight over proposed plans (Cain, 2012). Estimating the causal impact of these institutional reforms, however, is challenging for three reasons. First, as in any observational study, states that adopt redistricting reforms may differ in important ways from states that do not. Second, the reforms themselves intervene on complex and interrelated institutional processes, which differ widely even among the reforming states. Third, the outcome of the redistricting process is a districting plan, whose partisan features may be confounded by other factors such as a state's geography and demographics that are possibly correlated with reform efforts. Thus, both the treatment and outcome variables of our study are multifaceted and multidimensional. This means that most existing causal methodology, which considers a binary treatment and univariate outcome, is not readily applicable.

We estimate the causal effects of redistricting reforms by combining a formal theoretic approach with statistical methods of causal inference. To reduce the differing districting procedures across states to a single theoretically informed parameter, we model redistricting reform as a zero-sum sequential game. We collect the information about each state's relevant laws, starting from the initial map drawer (i.e., the redistricting commission or legislature) and following through varying stalemate processes and opportunities for court intervention. This new dataset is used to structure the players and available moves in the game. Analyzing the game allows us to estimate the ability of partisan players to maximize the partisan lean of a redistricting plan. The Nash equilibrium of the game is a measure of the "leeway" that a single party has over the final redistricting plan.

We then use these leeway measures as a continuous treatment variable in a differences-in-differences (DiD) design applied to the 2010 and 2020 redistricting cycles. This addresses potential confounding by comparing changes in states that have enacted reforms to those in similar states that have not, relying on a parallel trends assumption (Callaway et al., 2024). We estimate how changes in map drawer's leeway may influence resulting plans in terms of traditional measures of partisan fairness, such as the efficiency gap (Stephanopoulos and McGhee, 2015) and partisan symmetry (King and Browning, 1987; Katz et al., 2020), as well as nonpartisan "good-governance" outcomes, like the responsiveness to swings in partisan preferences.

Lastly, we separate the causal effect of state-specific institutional reforms from the change in political geography over time. To do this, we generate a sample of alternative redistricting plans via a simulation algorithm (McCartan and Imai, 2023) following each state's rules and political geography, but without regard to any partisan information (McCartan et al., 2022; Arora et al., 2024). We use these simulated plans as a nonpartisan baseline for both 2010 and 2020 redistricting cycles, and compute the difference in outcome variables between the enacted and simulated plans. This adjusts for state-specific changes in political geography, making the parallel trends assumption more plausible.

We find that more restrictive redistricting processes reduce partisan bias by constraining map drawers. For example, changing from a single-party legislature to an independent commission leads to about an reduction in excess seats of about 0.8 seats. We find smaller but positive effects of constraining reforms on electoral responsiveness. We estimate that a similar change from legislature to commission would increase the share of competitive seats in an average state from 25% to 32%.

One advantage of our methodological approach is the ability to conduct a counterfactual analysis of institutional reforms. We investigate how enacting recent procedural reforms nationwide could reduce widespread partisan gerrymandering. We quantify how the partisan bias and responsiveness of adopted plans could counterfactually change if all states nationwide adopted three kinds of commission structures currently enacted in several states: (1) an Ohio-style approach that requires supermajorities and uses a bipartisan backup commission, (2) a New York-style commission with a nonpartisan map drawer but several partisan veto points later in the process, and (3) a Michigan-style commission with a nonpartisan commission drawer, no partisan veto points, and the potential for court review.

We find that commissions can generally reduce the existing Republican bias but the details of commission structure matter. In particular, unlike reforms adopted in Ohio and New York, a Michigan-style non-partisan commission has no partisan veto points. We find that implementing Michigan-style reforms nationwide leads to an additional 22 Democratic seats, on average. Michigan-style reforms also increase electoral responsiveness whereas those adopted in New York and Ohio do not.

Beyond the substantive findings, this paper advances a new methodology for studying redistricting reform, one which we hope will translate to the study of other institutional systems as well. Redistricting reform efforts have produced diverse institutional changes across states. While most scholars have simply classified reforms into different categories (Cain, 2012; Edwards et al., 2017; Warshaw et al., 2022; Nelson, 2023), such an approach may miss important differences in these institutional changes.

In contrast, we use formal modeling to place these complex institutional characteristics on a continuous univariate scale to summarize how they constrain partisan actors. This theoretically driven approach, which we demonstrate predicts the empirical patterns well, makes it possible to apply the difference-in-differences strategy to our complex setting for credible causal inference.

While our work estimates the causal effects of redistricting reform *as a whole*, much of existing causal analysis has focused on one part of the process, such as registration patterns after the *Shelby* decision (Komisarchik and White, 2021) and turnout in packed districts (Fraga et al., 2022). Other studies have reported positive associations between the adoption of commissions and competitiveness (Carson and Crespin, 2004; Carson et al., 2014; Best et al., 2021; Nelson, 2023) as well as a negative association between commissions and incumbent protection (Nelson, 2023). These studies, however, do not adjust for various confounding factors considered in our analysis.

Beyond redistricting, our methodological approach can be seen as a general strategy for causal analysis of complex institutional reforms. By combining a game-theoretic treatment model with causal inference methods, we are able to leverage the strengths of these two approaches. A game-theoretic approach is a substantively effective way to model formal institutions with specific rules, making it possible to map multidimensional policies to a univariate summary in a theoretically-informed way. Once this summarization of various institutional reforms is done, we can apply standard causal inference methods to estimate the effects of counterfactual policies by mapping them directly to this univariate treatment variable.

Our work relates to a broader methodological literature in the social sciences. In economics, for example, Chetty (2009) advocates the combined use of structural and reduced-form approaches via a sufficient statistic in the manner similar to our analysis. In sociology, Lundberg et al. (2021) emphasize the importance of theoretically informed quantities of interest in causal analysis. Finally, in political science, Canen and Ramsay (2023) call for the integration of rigorous theoretical and empirical approaches in causal research. We demonstrate how such an analysis can be done in the estimation of causal effects of redistricting reforms.

2 Overview of Redistricting Processes and Reforms

Before introducing our methodological framework, we provide a brief overview of redistricting processes and reforms. Every decade following the U.S. Census, states and localities use many different procedures to redraw their legislative district boundaries. For example, states differ on whether legislatures or independent

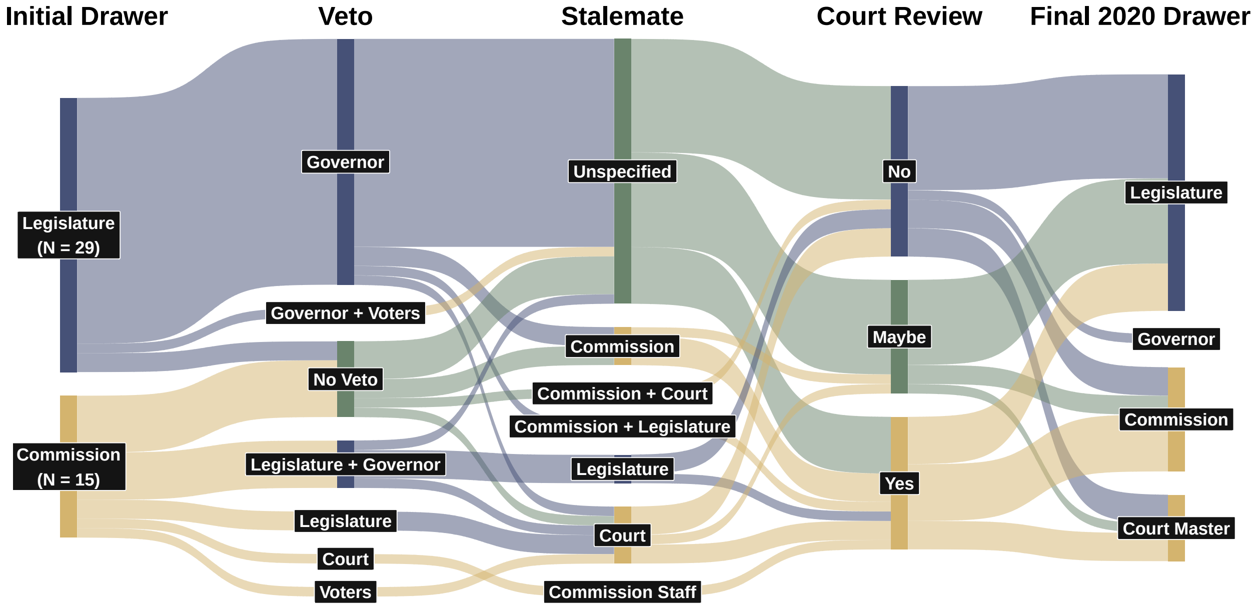

Figure 1: Redistricting procedures for all 44 states with more than one district in 2020. Each vertical column indicates a separate step in the redistricting process, and nodes indicate different values that each state can adopt at that step. The width of each area connecting the nodes is proportional to the number of states with that specific combination of procedure at both ends. Yellow nodes indicate actors or institutions that are not explicitly partisan, while blue nodes indicate explicitly partisan actors or choices. Here, we collapse multiple potential stalemate and veto procedures into one step for visual clarity (e.g., Governor + Voters indicates the possibility of a first veto by a governor, and a second by the voters).

redistricting commissions propose initial plans for new Congressional district maps. If these actors fail to produce a plan, state laws further vary on how these stalemates are handled. Some states pass responsibilities to a court (e.g., Virginia), a backup commission (e.g., Ohio), or a group of state party leaders (e.g., Iowa). Even plans that pass the proposal step can face a veto from other actors like the governor or state legislature.

As explained in Section 3.1 below, we collect information about the redistricting process used in each state for the 2010 and 2020 cycles. Figure 1 summarizes this dataset and illustrates the diversity of 2020 redistricting procedures across states. For example, 29 states drew initial plans in their legislature, while 15 states used an independent redistricting commission instead. There is significant variation in procedures following this initial draw, with seven distinct veto mechanisms across all the states. Much of this variation originates from state-specific policy campaigns that have adopted widely different goals driven by local actors (Keena et al., 2021). For example, California created a redistricting commission that has the power of drawing congressional district boundaries through a series of ballot propositions in the early 2000s. In recent years, many states have also implemented redistricting reforms often driven by desires to limit partisan influence. For example, Michigan voters approved a constitutional amendment to establish an independent redistricting commission in 2018.

These procedural differences make it difficult to attribute causal effects to particular institutional designs or reforms. Most existing work has turned to classification schemes that simplify this variation by assigning the control of map redrawing to a single drawer, typically the creator of the initial or final plan (Edwards et al., 2017; Carson et al., 2014). A more complex alternative would be to create indices that count the number of times a particular action appears in the state procedure (e.g., the number of veto points). For example, as shown in Figure 1, even states that have a commission propose their initial plan vary drastically in how the process continues afterwards.

While these simplifications make the study of complex procedures tractable, they risk oversimplifying complex procedures in two ways. First, attributing institutional outcomes to a single actor ignores the fact that the final redistricting plan is produced through a series of steps. Take, for example, the New York redistricting process in 2020. An independent bipartisan commission had the power to propose an initial plan, but failed to agree on a single plan. State law required the stalemated process to move to the legislature, which adopted a plan that Republican and civil rights groups criticized as a Democratic gerrymander. These groups challenged the adopted plan in a series of lawsuits, and in 2022 the New York Court of Appeals struck down the plan and tasked a court-appointed special master with drawing a remedial plan. Though the final plan scores well on quantitative fairness metrics (Kenny et al., 2023), simply classifying a "court" as the sole map drawer for New York in 2020 overlooks the partisan interests involved, failing to capture the complexity of the procedures that led to the final plan.

Second, simple classification schemes can overlook strategic interactions, where the behavior of certain actors depend on the presence or characteristics of others. For example, a commission might draw a different map if it knew the map could later be reviewed by a court. Or, the potential of a governor's veto may limit the likelihood of partisan gerrymander by the state legislature, but not in cases where the governor and legislative majority share a partisan affiliation. More detailed coding of procedural schemes can account for these interaction effects and increase realism, but will necessarily decrease statistical power, making it more difficult to estimate causal effects.

Thus, we face a methodological dilemma: while some simplification of institutional features is necessary, common approaches to doing so obscure critical characteristics of redistricting processes. In the next section, we propose a theoretically-grounded approach that models redistricting as a zero-sum sequential game and uses its Nash equilibrium as a treatment variable summarizing complex institutions.

3 A Theoretical Model of Institutional Leeway

To study the impact of redistricting reforms and processes, we first develop a standardized set of 14 institutional features that capture the most important actors in the congressional redistricting processes of the states. We collect the data on these features so that it is possible to compare redistricting processes consistently across states, despite the differences in laws and bodies governing redistricting in each state.

Next, we develop a sequential redistricting game to summarize these features in a theoretically informed manner. Specifically, we use the Nash equilibrium of this game as a one-dimensional measure of the institutional leeway political actors have over the partisan lean of the resulting redistricting plan. Finally, we empirically validate the proposed measure by demonstrating that it is not particularly sensitive to model specification and predicts redistricting outcomes well.

3.1 Data on the relevant institutional features

Our standardized coding of relevant institutional features is based on the prototypical redistricting process shown in Figure 2. First, an initial map drawer proposes a plan, which may be vetoed by other actors. If the plan is not vetoed, it can be challenged in court or not. If it is vetoed, there is another round of map drawing. If a plan is vetoed twice, or if the initial map drawer cannot agree on a plan, then a different institution (often a court) must resolve the stalemate and adopt a plan.

Every state's process can be described as a subset of this prototype. For example, in Alabama, the legislature draws congressional districts, subject to a governor's veto, and there is no mechanism for state court review. Thus, Alabama's process would be described by the "Drawer" and "Veto 1" steps only in Round 1 of the Figure 2 process.

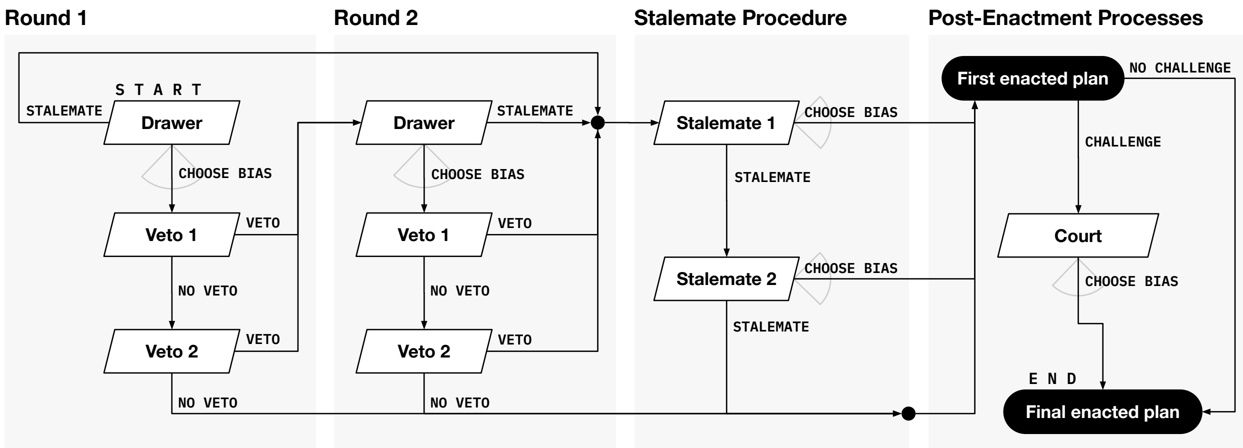

Figure 2: Prototypical game tree used to model redistricting in every state. States differ in which party, if any, controls each node, and which nodes are present in the state's process.

For each state, we record which institutional body, if any, acts at each step, and which party, if any, controls that institution. We also collected additional information relevant to modeling redistricting processes and court review, such as whether the state's redistricting plans were subject to DOJ preclearance before 2013,[2] and which institutional body ended up drawing the plan that was used in the first postcensal elections. Most procedural details are straightforward, and the information for coding is readily available in public data from each state.[3] Furthermore, many states have similar redistricting procedures that are easily classified under the categories we defined earlier.

Appendix A explains in detail how each of these variables was coded and describes special cases. Figure 1 of the previous section graphically summarizes this dataset for 2020 while Table A2 of the Appendix presents all the variables across states in both 2010 and 2020.

3.2 The redistricting game

Coding the details of each state's redistricting process preserves important procedural information, compared to simply categorizing each state into a small number of groups such as "legislature-controlled" and "independent commission." However, the detailed coding presents a challenge for causal inference, since the treatment—a state's redistricting process—is now high-dimensional. Out of the 87 state-decade processes we code,[4] there are 58 distinct combinations of procedural variables, 41 of which are completely unique.

The combination of high-dimensional treatment and limited sample size means that there may not be enough information to estimate the causal effect of changing from one specific configuration of procedural variables to another, without further assumptions.

We address this methodological challenge by leveraging two basic substantive assumptions about the redistricting process: first, that each party aims to draw a map that favors it as much as possible, and second, that the parties are constrained by statutory and constitutional rules in doing so. Specifically, we treat the redistricting process depicted in Figure 2 as a sequential zero-sum game[5] with two players—the Democratic and Republican parties—each trying to maximize the degree to which the drawn plan favors their party.

[2]The *Shelby* decision was handed down in 2013, ending DOJ preclearance.
[3]For instance, from https://redistricting.lls.edu/.
[4]Only 43 states had two or more congressional districts in the 2010 cycle; for 2020, Montana did as well, bringing the total to 44.
[5]The zero-sum assumption rules out cooperation between the parties to, e.g., protect incumbents. Given the role of the game here as a data summary rather than predictive model, we believe this is an acceptable limitation.

In each state, the nodes in the game tree can be controlled by different parties, or by neither party, e.g., when a supermajority is required to adopt a plan and neither party has supermajority control. Any split-control nodes, as well as a node for state court action, are considered moves by nature. Some nodes involve discrete choices such as whether to veto a plan or not, others are labeled “choose bias,” meaning that the player at that node draws a plan with a chosen amount of partisan bias favoring either party. This bias is exactly the utility to the party of the chosen plan: a plan with bias x is worth x to the Republicans and $-x$ to the Democrats. We need not quantify exactly what the utility measure is as a function of a specific plan chosen; it suffices to let the parties try to maximize an abstract univariate measure of partisan bias. We let the bias score range from −4, indicating a maximum Democratic advantage, to +4, indicating a maximum Republican advantage.

For concreteness, consider the case of redistricting in Alabama, as described earlier. The first move belongs to the Republican party, which controls the state legislature. It has to pick the amount of partisan bias $-4 < x < 4$ in the plan that it adopts. If this plan is ultimately the final enacted plan, the Republicans receive utility x and the Democrats receive utility $-x$. The second move also belongs to the Republican party, which controls the governorship. The party must decide whether or not to veto the adopted plan; if it does, there is a second round of drawing and vetoing by the legislature and governor (controlled by Republicans).

If a plan is adopted at either the first or second round, it proceeds to possible court review. The court may decide to accept a legal challenge, decide in favor of the plaintiffs, and redraw the map; this choice is considered a move by nature. If the court review results in a redrawn map with partisan bias x', then the Republicans receive utility x' and the Democrats receive utility $-x'$. In Alabama, court review is generally not allowed on partisan grounds, but challenges under the federal Voting Rights Act are possible, so there is moderate probability of the legislature-adopted plan being overturned. If the legislature’s plan is vetoed both times, the process results in a stalemate. Since there is no enumerated stalemate procedure in Alabama, courts must step in and redraw district lines to ensure compliance with federal constitutional “one person, one vote” apportionment requirements. This is also considered a move by nature.

To complete the description of the game, we must specify the rules for determining expected outcomes for moves by nature. There are two kinds of moves by nature: map drawers controlled by neither party exclusively, and the results of court challenges. The full specification of these moves may be found in Appendix B, but we briefly summarize them here. We make three assumptions: (1) nonpartisan map-drawers whose choices are subject to veto will generate maps which favor the party controlling the veto after a first veto has been made, (2) stalemate map-drawing will produce a map that is moderately balanced but tends to be influenced by any biases present in the most-recent redistricting proposal, and (3) split-control map-drawers will stalemate with some probability and produce similar results to nonpartisan map-drawers the rest of the time.

Finally, we decompose the court challenge process into five components: the probability that a legal challenge is possible, the probability that a challenge is made when possible, the probability that a court sides with plaintiffs, the expected remedy a court orders in those cases, and the probability and expected effect of a challenge based on the federal VRA. The latter applies only to states previously subject to DOJ preclearance. Each of these five components has a parametric specification, detailed in Appendix B. The partisan control of state courts is accounted for in the specification of these various components, thus allowing judicial polarization to enter the picture without assuming absolute strategic coordination between the state party and its allies in the judiciary.

All in all, the game specification depends on 19 parameters which govern the moves by nature, with most of these parameters relating to the court challenge process. Rather than fix these parameters to arbitrary constants, we place a prior distribution on each parameter over a range of probable values. We simulate 100 different draws from this joint prior distribution; each draw generates a slightly different game specification. We then average our results across the random draws.

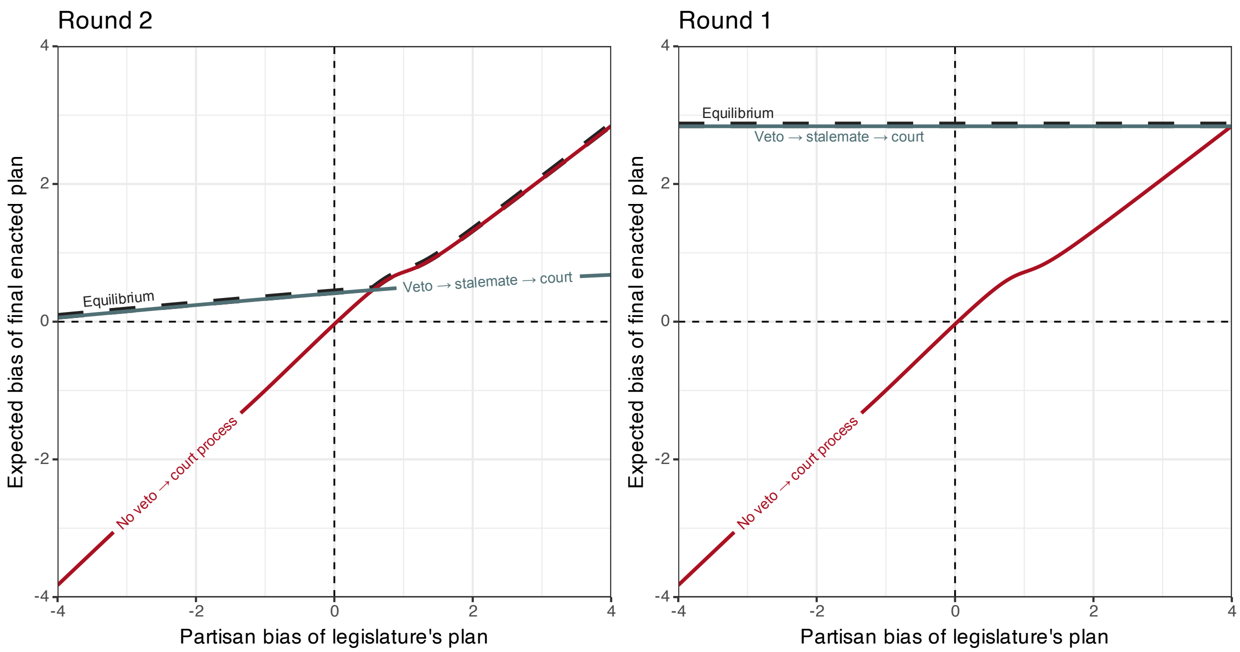

Figure 3: Expected outcomes for "veto" and "no veto" moves by Alabama's Republican governor in Round 2 (left) and Round 1 (right), as a function of the partisan bias of the legislature's proposed plan. The equilibrium outcome at each node is indicated by the dotted line.

3.3 Equilibrium solution as a treatment variable

To solve the problem of a high-dimensional treatment, we use the subgame perfect Nash equilibrium of the redistricting game. The equilibrium solution captures the expected partisan bias of plans that arise out of a state's redistricting process, under the current party control of the state's institutions. All of the multi-step institutional interactions and negotiations that might happen as part of the redistricting process are thus reduced to a univariate score. The upshot is that the 14 procedural variables can be reduced into a single axis that measures the leeway political actors have over the partisan lean of the final redistricting plan.

To calculate the equilibrium itself, we numerically solve the game via backwards induction. This requires up to four levels of nested optimization. For Alabama, this means that we start with the last partisan move in the game, which is the Round 2, Veto 1 player, the Republican governor. She must decide, given a proposed partisan bias x from the Round 2 map drawer, whether to veto the plan, in which case the enacted plan is determined via the court stalemate process, or to accept the plan and let it become law. In either case, the resulting plan may be challenged in court.

The expected outcome of both the stalemate-then-court and court-alone choices are determined by the parameters governing the moves by nature. They are shown as a function of the bias of the legislature's proposed plan in Figure 3(a), for a typical set of game parameters. In this case, because the only realistic avenue for court review in Alabama is a VRA claim, the expected outcome of a court challenge given an enacted plan x is roughly x. The expected bias of a plan drawn through stalemate procedures under a Republican court is roughly 0.5, though it also increases monotonically in x. Since the governor wishes to maximize the signed partisan bias, she will veto the plan if its bias is less than about 0.5, and not veto it otherwise.

Having solved the Round 2, Veto 1 subgame, we move backwards to the Round 2 drawing subgame. The Republican legislature must decide on the partisan bias of the plan it proposes, or can choose to stalemate by failing to adopt a plan. If it stalemates, then the expected outcome will be given by the gray "Veto" line in

Figure 3(a), and will depend on the partisan bias of the plan originally adopted by the legislature in Round 1 before being vetoed. If it does not stalemate, then the expected outcome is determined by the dashed equilibrium line in Figure 3(a). To maximize the partisan bias, the legislature at Round 2 will draw a map with bias of +4, leading to expected bias (utility) of 2.84 for this set of game parameters.

We can then move backwards again and consider the Round 1, Veto 1 subgame. Figure 3(b) shows the expected outcomes under each choice, just as before. In equilibrium, the governor will veto any plan with bias less than 4; regardless of the governor's decision, the expected bias is 2.84. Finally, the legislature can make any initial proposal, since the expected outcome is the same regardless. Thus, the overall equilibrium is a bias of 2.84. This solution process is automated across all of the states and is carried out on each of the 100 different draws from the prior on the game's parameters.

For our causal analysis, our first treatment measure is exactly the game's Nash equilibrium, averaged across the prior. This measure depends on which parties control each node in the game tree. We also generate a second treatment measure which does not depend on current party control. This is calculated by assigning a single party (here, the Democrats[6]) to control each node in the game tree that belongs to a partisan actor—legislature, governor, or partisan commission—and then recalculating the average Nash equilibrium. We refer to this equilibrium as the *maximum leeway* of a state's redistricting process, since it captures the expected bias under a worst-case partisan outcome where all of the levers of state government are controlled by one party. The first treatment measure we refer to as the *realized leeway*, since it depends on the realized values of party control for each state institution.

Figure 4 visualizes these treatment measures for all the states we study; an arrow indicates a change in the treatment value from 2010 to 2020. The realized leeway scores (game equilibria) cover the entire range of possible partisan biases, from West Virginia in 2010 with complete Democratic control of state government, to Wisconsin, Iowa, Kansas, Utah, and Nebraska, where Republican trifectas were unconstrained by Democrats or by the VRA in both cycles. The maximum leeway scores likewise span the range of possible bias, but take on fewer values, since the various actual combinations of partisan control of state institutions are no longer considered.

As might be expected, states with independent redistricting commissions, such as Michigan, have a realized and maximum leeway of zero. In 2010, Michigan redistricting was entirely controlled by the legislature and governor, who were both Republican. By 2020, the process was reformed to have a strong independent commission. In the new process, the commission is bound by strict criteria, the process allows for court review, and even the commissioners themselves are selected in part by a lottery system. With the reform, Michigan's maximum leeway changed from the highest observed value to zero, meaning that partisan actors have no control over redistricting outcomes, on average.

Other states' scores are similarly interpretable, where introducing a commission (e.g. N.Y., Va., or Colo.) drastically reduces its leeway to match states with similar sets of rules in 2020. States that had intervening litigation to clarify the interpretation of state redistricting rules (e.g. N.M., N.C., or Pa.) see appropriate changes in leeway, even absent a commission. Further, in our party-signed treatment, states that had a flip in party control of the whole system see very large changes in leeway in the correct direction (e.g. W.Va.). Finally, states with minor changes to the total control of state institutions see minor changes in leeway (e.g. N.C.).

[6]The game is completely symmetric between parties, with one exception: due to racial polarization in the U.S., challenges to redistricting plans under the VRA are much more likely for Republican-leaning plans, and when these challenges prevail the remedy almost always results in a plan more favorable to Democrats. In keeping with the goal of calculating the *maximum* leeway in each state, we therefore assign states to Democratic control under the counterfactual scenario here. Assigning to Republican control instead would slightly reduce the calculated leeway for those states previously subject to DOJ preclearance.

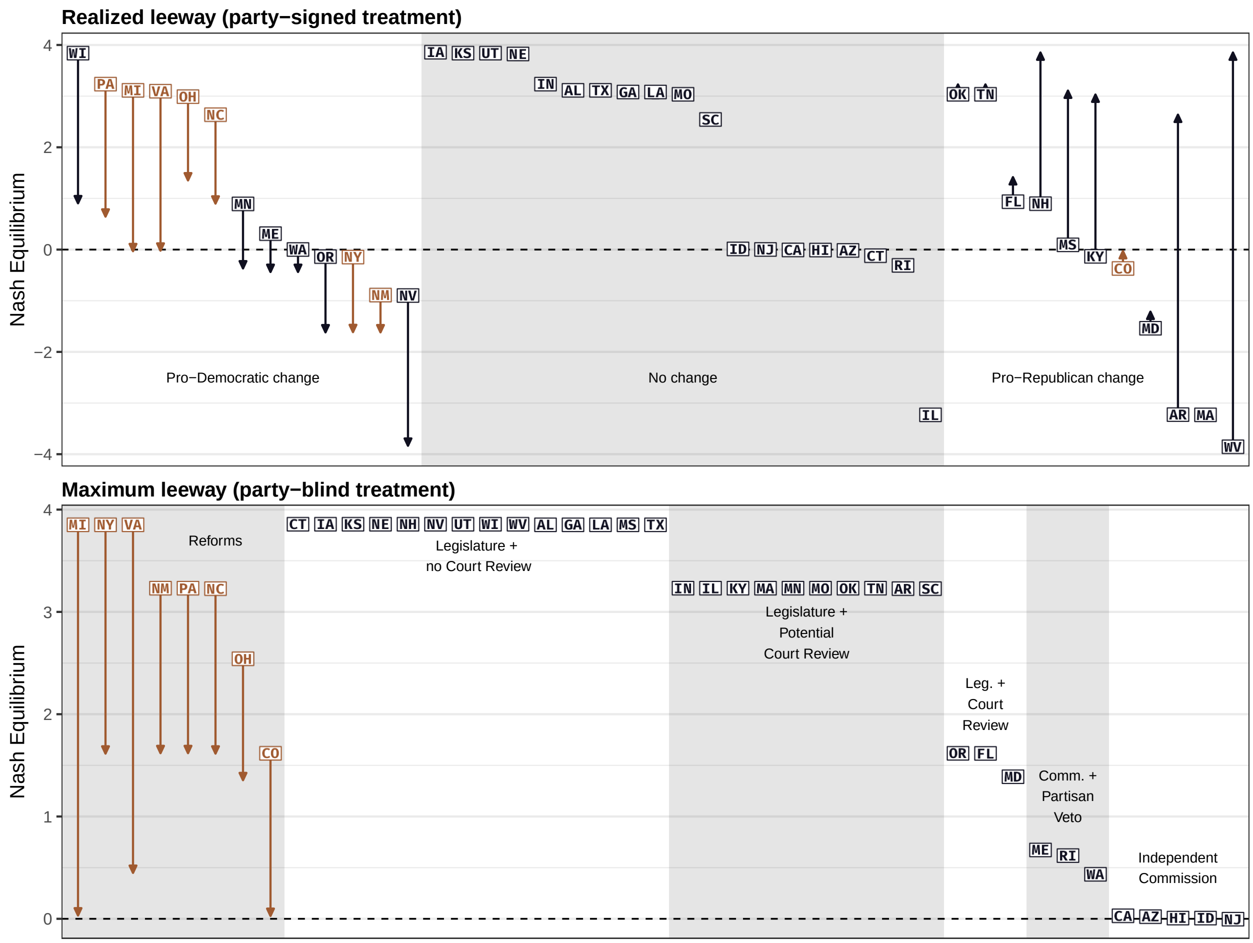

Figure 4: Treatment values for each state in 2010, with values for 2020 indicated by arrows, where different. States in orange are those which experienced a reform to their redistricting procedures, either by legislation, constitutional amendment, or a court ruling that allowed for state court review of alleged partisan gerrymanders.

3.4 Validation

Building treatment variables through a game-theoretic model provides significant advantages in dimension reduction and interpretability, but it also comes with the risk that the model is misspecified. Although it is difficult to completely verify the validity of each component of the model, we take several steps towards empirically validating the model as a whole, showing that the resulting treatment variable predicts the observed outcomes well.

The model was designed to be flexible enough to capture the actual redistricting processes in each state. Any misspecification is therefore due to how the moves by nature are specified. First, we find that the treatment values are not sensitive to specific parameter values, up to monotonic transformations. Across the 100 random draws from the prior, the average pairwise Spearman correlation between the Nash equilibria for each state is above 0.99.[7] This also gives us confidence that the specific choice of prior is not influencing the results.

Second, as shown in Figure 5, there is a meaningful correlation in the expected direction between the treatment values and measures of partisan advantage. The left panel plots the expected share of seats won by Republicans

[7] i.e., for each pair of random draws, we calculate the Spearman correlation between the two vectors of calculated state equilibria, then average these correlations across all pairings.

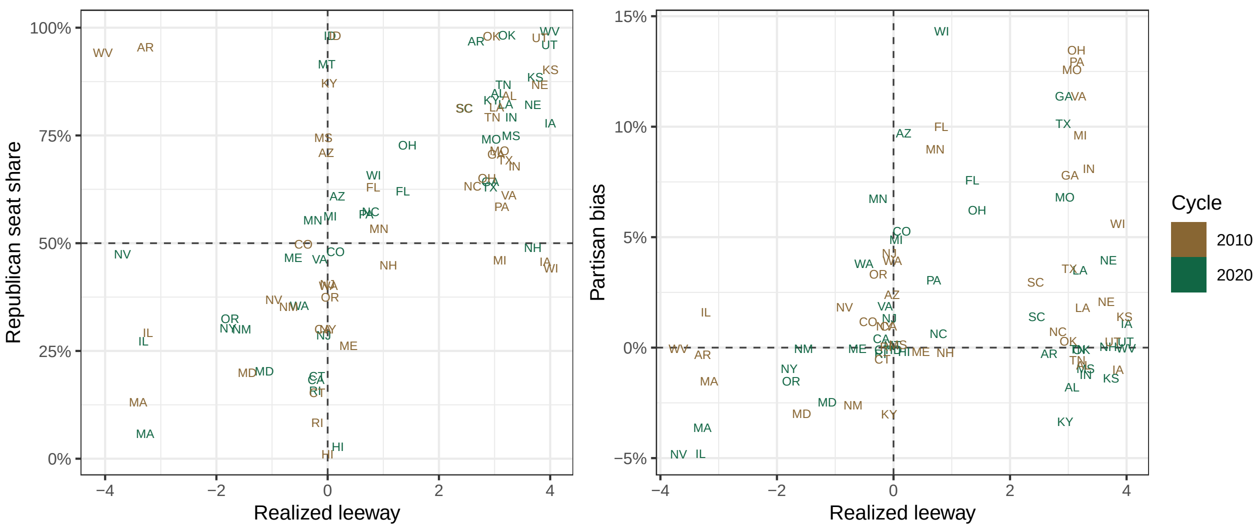

Figure 5: Measures of partisan advantage versus treatment values for states' enacted plans for the 2010 and 2020 redistricting cycles. Points are slightly jittered to avoid overplotting.

Table 1: Correspondence between the institution that drew the final redistricting plan for each state and the most likely outcome based on the equilibrium path of the redistricting game.

	Most likely in equilibrium			
Final drawer	Legislature	Commission	Court	Total
Legislature	31.9	0.0	19.1	51
Commission	0.0	18.9	3.0	22
Court	1.7	0.0	12.3	14
Total	33.6	18.9	34.4	87

versus the realized leeway measure; recall that the leeway measure is positive for plans that favor Republicans. The right panel shows the partisan bias (King and Browning, 1987), measured at the state's baseline vote share, versus realized leeway. Larger values of partisan bias correspond to plans that systematically favor one party, with positive values favoring Republicans.

For both outcome measures, the purely *a priori* model predictions of the expected partisan bias do in fact correlate with the actual partisan bias of the plans that come out of each state's redistricting process. To what extent these correlations can tell us about the causal effects of reforms is of course another matter, and the primary question at issue in this paper.

We also compare the equilibrium path in the model with the actual institution that drew the final map in each state. Because the moves by nature in the game are random, the actual equilibrium path may be a probabilistic mixture over multiple possible paths. We average these probabilities over the 100 draws of the game to arrive at an overall probability that a legislature, commission, or court draws the final map in each state, given its procedure and party control.

Table 1 compares the most likely final map-drawer, according to these probabilities, to the actual institution that drew the final redistricting plan. The correspondence is generally excellent, considering the predictions are made based on theory alone. There is a tendency, however, for the model to predict court-drawn maps when they are in fact drawn by legislatures. The overall likelihood of a court intervention is itself a model parameter that is varied across random draws, and as noted above the ranking of the states by leeway is basically unchanged by different parameter values. Therefore, we expect this tendency to over-predict court intervention to at most slightly understate the equilibrium value for states with court review, since a higher likelihood of intervention would tend to pull the equilibrium towards zero.

Finally, it is natural to wonder whether a more accurate model could be obtained by estimating the model parameters from the observed data rather than specifying their prior distributions. That is, one could find the parameter values that maximize the correlation between the treatment and outcome values. We do not take this approach in order to cleanly separate out estimation of causal effects from construction of a treatment variable. This enables us to employ a causal identification strategy that does not assume the correct specification of the game theoretic model. Instead, we use the model to summarize a high-dimensional treatment variable in a theoretically informed way without looking at the outcome variable.[8]

4 Estimating the Causal Effects of Institutional Leeway

To estimate the causal effects of changes in leeway on redistricting outcomes, we use a differences-in-differences (DiD) design with a continuous treatment variable. This strategy addresses potential confounding by comparing changes in states that have enacted reforms to changes in similar states that have not modified their redistricting process across the 2010 and 2020 redistricting cycles. We assume that, in the absence of reforms, the states with institutional changes would have experienced the same trend in outcome variables as those states without reforms.

Formally, let $\mathbf{Z}_{it}$ be the 14-dimensional vector of institutional features for state i at time t discussed in Section 3.1. We use $t = 0$ and $t = 1$ to denote 2010 and 2020 redistricting cycles, respectively. Let $u^*(\mathbf{z})$ represent the equilibrium outcome (game utility) for the average Nash equilibrium of the redistricting game described by the process $\mathbf{z}$. Then, we can define a univariate treatment variable $D_{it} = u^*(\mathbf{Z}_{it})$ for each state-decade, which represents the "dose" of institutional leeway given to political actors.

The key assumption made by our use of the Nash equilibria as the treatment variable is that the institutional features affect the outcome only through this treatment variable, i.e.,

$$Y_{it}(\mathbf{z}) = Y_{it}(\mathbf{z}') \quad \text{for any} \quad \mathbf{z}, \mathbf{z}' \text{ with } u^*(\mathbf{z}) = u^*(\mathbf{z}'),$$

where $Y_{it}(\mathbf{z})$ denotes a generic potential outcome variable for state i at time t with institutional features $\mathbf{z}$.[9] Under this assumption, we can simply write the potential outcomes as $Y_{it}(d) = Y_{it}(u^*(\mathbf{z}))$ for the treatment dose $d = u^*(\mathbf{z})$.

Our target estimand is the conditional average treatment effect (CATE) for a change from treatment level d to d' given a set of covariates $\mathbf{X}_i$ that are not affected by the treatment. The CATE is defined as,

$$\text{CATE}_{\mathbf{x}}(d', d) = \mathbb{E}[Y_{i1}(d') - Y_{i1}(d) \mid \mathbf{X}_i = \mathbf{x}].$$

Estimating $\text{CATE}_{\mathbf{x}}(d', d)$ for any pair of values d', d and covariate values $\mathbf{x}$ requires the identification of a full dose-response curve. Our identification approach relies on a strong conditional parallel trends assumption introduced by Callaway et al. (2024). Specifically, we assume that the average change in outcomes for states experiencing a change in treatment dosage from d to d' depends only on the observed covariates, and not the observed dosage value.

[8] We did attempt to fit the game parameters to the observed data for the 2010 cycle alone via maximum likelihood, but we found that the observed data were only minimally informative about most of the parameter values. This is good news, insofar as it implies that a wide range of parameter values are all compatible with the observed data.

[9] For identification, it in fact suffices to assume only that $\mathbb{E}[Y_{it}(\mathbf{z}) \mid \mathbf{X} = \mathbf{x}] = \mathbb{E}[Y_{it}(\mathbf{z}') \mid \mathbf{X} = \mathbf{x}]$ for any $\mathbf{z}, \mathbf{z}'$ with $u^*(\mathbf{z}) = u^*(\mathbf{z}')$; however, the stronger individual-level assumption is invariant under transformations and ensures that no additional assumptions are required for estimation.

Our covariates include a range of variables that existing work in the redistricting literature identifies as predicting redistricting outcomes and are not affected by the redistricting reforms. We include (pre-treatment) 2008 Democratic presidential vote share, an indicator for being in the South,[10] the logarithm of the number of districts in 2020, the change in the number of districts between 2010 and 2020, the median vote share to candidates not from the Democratic or Republican parties between 2000 and 2010, the logarithm of the average number of state corruption convictions by year between 2000 and 2010, and an indicator for whether states allow ballot initiatives.

Formally, we require the following conditional parallel trend assumption for all d, d':

$$\mathbb{E}[Y_{i1}(d') - Y_{i0}(d) \mid \mathbf{X}_i = \mathbf{x}] = \mathbb{E}[Y_{i1}(d') - Y_{i0}(d) \mid \mathbf{X}_i = \mathbf{x}, D_{i0} = d, D_{i1} = d'].$$

This assumption is analogous to the parallel trends assumption in the traditional binary difference-in-differences design, but differs in that it refers to changes in dosage across a continuous measure, rather than a single level change from zero. This conditional parallel trends assumption allows us to identify our estimand as

$$\text{CATE}_x(d, d') = \mathbb{E}[Y_{i1} - Y_{i0} \mid \mathbf{X}_i = \mathbf{x}, D_{i0} = d, D_{i1} = d'] - \mathbb{E}[Y_{i1} - Y_{i0} \mid \mathbf{X}_i = \mathbf{x}, D_{i0} = d, D_{i1} = d],$$

where $Y_{it} = Y_{it}(D_{it})$ is the observed outcome. Averaging over the marginal distribution of $\mathbf{X}_i$, we can also estimate the average treatment effect (ATE).

To further increase the credibility of the conditional parallel trend assumption above, we adjust for the changes in political geography between the two redistricting cycles. Partisan preferences and the geographic distribution of voters can change in different ways within each state over time. For example, Michigan, which experienced reform, became more Republican during our sample period, while Georgia, which did not, became more Democratic. Without this adjustment, an estimated effect of the Michigan reform would be biased by the differential change in the states' political geographies.

To make the adjustment, we use representative sets of nonpartisan redistricting plans within each state that respect each state's specific redistricting rules. These simulation samples were separately generated for 2020 (McCartan et al., 2022) and 2010 (Arora et al., 2024) using the algorithm of McCartan and Imai (2023). We subtract the mean outcome in each simulated sample, denoted by $\widetilde{Y}_{it}$, from the observed outcome Y_{it} before estimating the causal effects under the DiD design. We argue that this subtraction of the outcomes based on simulated baseline plans from the observed outcomes accounts for the change in political geography in each state.

Thus, we assume that once we adjust for the state-specific change in political geography, states with different changes in institutional features have the parallel trend in the potential outcome. Formally, if we denote the difference between the potential outcome and the simulated outcome as $\Delta Y_{it}(d) = Y_{it}(d) - \widetilde{Y}_{it}$, our conditional parallel trend assumption becomes

$$\mathbb{E}[\Delta Y_{i1}(d') - \Delta Y_{i0}(d) \mid \mathbf{X}_i = \mathbf{x}] = \mathbb{E}[\Delta Y_{i1}(d') - \Delta Y_{i0}(d) \mid \mathbf{X}_i = \mathbf{x}, D_{i0} = d, D_{i1} = d']$$

for all d, d'. Then, as above, the CATE is identified as

$$\text{CATE}_{\mathbf{x}}(d, d') = \mathbb{E}[\Delta Y_{i1} - \Delta Y_{i0} \mid \mathbf{X}_i = \mathbf{x}, D_{i0} = d, D_{i1} = d'] - \mathbb{E}[\Delta Y_{i1} - \Delta Y_{i0} \mid \mathbf{X}_i = \mathbf{x}, D_{i0} = d, D_{i1} = d],$$

where $\Delta Y_{it} = Y_{it} - \widetilde{Y}_{it}$.

[10] Southern states are defined as Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, and West Virginia. This set corresponds to Confederate states, plus Kentucky and West Virginia, which also had a history of Democratic dominance in the 20th century and notable differences between party preferences at the state and national levels.

We estimate this causal estimand with a Bayesian linear regression model, where the response is the change in each simulation-adjusted outcome between 2010 and 2020, $\Delta Y_{i1} - \Delta Y_{i0}$. Given the small sample size ($n = 87$) and the moderate number of covariates ($p = 15$ with interactions), a Bayesian model is useful, as the priors help avoid overfitting to a few samples, and uncertainty is automatically quantified. A more flexible regression model beyond linear would be unlikely to increase predictive power, given the small sample size, and would further risk overfitting. The predictors are the change in treatment level $D_{i1} - D_{i0}$, the baseline treatment level D_{i0}, the covariates $\mathbf{X}_i$, and the interaction of treatment change with baseline treatment and with each of the covariates.

In terms of the prior specification, the intercept receives a weakly informative prior of $\mathcal{N}\left(0, (2.5 \cdot \sigma_y)^2\right)$, where σ_y^2 is the variance of ΔY_t. Each main effect coefficient receives a weakly informative prior of $\mathcal{N}\left(0, (0.75 \cdot \sigma_y/\sigma_x)^2\right)$, where σ_x^2 is the variance of each covariate x. The prior standard deviation is further scaled by 0.25 for interaction terms, following the discussion in Gelman et al. (2008). Finally, the residual standard deviation receives an $\text{Expo}(1/\sigma_y)$ prior. Computation is carried out via NUTS-HMC as implemented in the Stan probabilistic programming language (Stan Development Team, 2024).

5 Estimated Causal Effects of Redistricting Reforms

The outcome of the redistricting process is a complete congressional districting plan. A plan can be evaluated in a number of ways—how many seats each party is expected to win, how many of the seats are competitive, how well the partisan composition of the delegation matches the voters' preferences, and so on. We begin this section by introducing quantitative measures of several partisan and nonpartisan aspects of districting plans. These constitute the outcome variables for the causal estimates we present in the second half of the section.

5.1 Outcome measures

Our measures can be divided into two buckets: nonpartisan outcomes that quantify how much gerrymandering is present or how responsive a districting plan is to shifts in public opinion, and partisan outcomes that capture which party is advantaged by a districting plan.

All of these measures are calculated from district-level election results. Since there is significant exogenous variation in election results due to swings in the national political environment, rather than using actual 2012 and 2022 House election results to evaluate districting plans, we use a statistical model to estimate the distribution of election results across future hypothetical elections. A statistical election model also allows us to apply these same measures to the simulated districting plans, under which no elections have taken place.

We adopt the model from Kenny et al. (2023), which assumes district election outcomes can be decomposed into a baseline district-level vote share plus district-specific and national swings. The model is closely related to the stochastic uniform partisan swing model of Gelman and King (1994) and the congressional model of Ebanks et al. (2023). As the baseline district-level vote, we use the 2008 presidential results for the 2010 redistricting cycle and an average of the 2016 and 2020 presidential results for the 2020 cycle, each logit-shifted so that the national partisan vote is exactly 50/50. As in Kenny et al. (2023), we use historical congressional election data since 1976 to estimate the variance of the district and national swings. We calculate exact expectations of all of our outcome measures against the predictive distribution of the model via numerical integration. The full details of this approach can be found in Appendix D.

Our primary partisan measure of districting plans is the expected number of seats won by the Republican party. As discussed in Section 4, we subtract the average outcomes in the simulated baseline redistricting plans (McCartan et al., 2022; Arora et al., 2024) from the observed outcome in the enacted plan to adjust for the state-specific change in political geography, which is a potential confounding factor. After this adjustment, the seat outcome ranges from –1.57 in Illinois in 2020 (favoring Democrats) to 1.78 in Texas in 2020 (favoring Republicans). These simulation-differenced seats can be directly interpreted as a measurement of bias due to partisan gerrymandering, with positive values indicating a Republican bias beyond what would be expected based on the state's political geography alone.

Republican seats is an interpretable measure, but may not be completely comparable across states. Depending on political geography and especially the total number of districts in each state, the natural variation in the number of seats outcome may vary significantly between states. To address this, we also include as an outcome measure the simulation z-score of Republican seats, which is calculated by taking the simulation-differenced seats outcome and dividing it by the standard deviation of Republican seats in the simulation set (Kenny et al., 2023). This puts all the states' outcomes on a common scale, increasing the plausibility of the parallel trends assumption and the homoskedasticity assumption of the estimation model.

For a nonpartisan outcome measure, we take the absolute value of these two partisan measures. Both the absolute difference and the absolute z-score of Republican seats measure how far the enacted plan deviates from the nonpartisan simulation baseline in terms of partisan composition.[11]

We also study two more outcome measures based on the efficiency gap, a popular measure of partisan bias that is calculated as the difference between the two parties' wasted votes (Stephanopoulos and McGhee, 2015). Wasted votes are votes cast for losing candidates and votes cast for winning candidates beyond the threshold needed to win. The idea behind the efficiency gap is that a fair plan should have similar numbers of wasted votes for both parties. The partisan-signed version of the efficiency gap indicates which party benefits from the plan in terms of vote efficiency; we use positive values to indicate pro-Republican bias. Unlike with Republican seat share, we do not subtract the average simulated value from the partisan-signed efficiency gap measure. This is because the efficiency gap is already designed to have a universal scale, where 0 indicates partisan fairness; subtracting the simulated average would change the definition of a fair baseline for this measure. To create a nonpartisan measure, as with Republican seats we take the absolute value of the difference between the enacted and average simulated plan's efficiency gap. Smaller values indicate a fairer plan by this measure (closer to the normatively preferred value of 0).

Finally, we also measure the responsiveness of districting plans to changes in the national electoral environment. Responsiveness is measured as the rate of change in the share of Republican seats given an infinitesimal change in Republican vote share nationwide. Responsiveness is closely linked to the presence of competitive seats: the more competitive seats there are, the larger the change in seat share will be for a given shift in vote share. Indeed, electoral responsiveness is the primary motivation for the creation of competitive congressional districts.

The responsiveness measure ranges from 0.044 in Wyoming in 2020 to 7.91 in New Hampshire in 2020, with most plans' values lying between 0.5 and 3. The interpretation of these values is as follows: in New Hampshire, a 1pp increase in a party's vote share leads to a 7.91pp change in the party's expected seat share. This relatively large increase makes sense in context—both of the state's congressional districts have Republican vote share within a few points of 50%. As with the other measures, we subtract the mean responsiveness of the simulated plans from the enacted plan's responsiveness when estimating causal effects.

In Appendix C, we extend the analysis to a series of alternative measures of partisan bias and find qualitatively similar results.

5.2 Results

We first study the effects of changes in maximum leeway on Republican seats. Recall that maximum leeway represents the worst expected bias under a single party control. The left panel of Figure 6 shows the estimated

[11] For the transformed outcomes based on z-score and absolute value, we do not interpret the estimated causal effects in terms of the original outcome variable (i.e., the expected number of Republican seats). Instead, the estimates should be interpreted as the effects on the transformed outcomes.
This also means that the parallel trends assumption is assumed to hold on the transformed outcome variable.

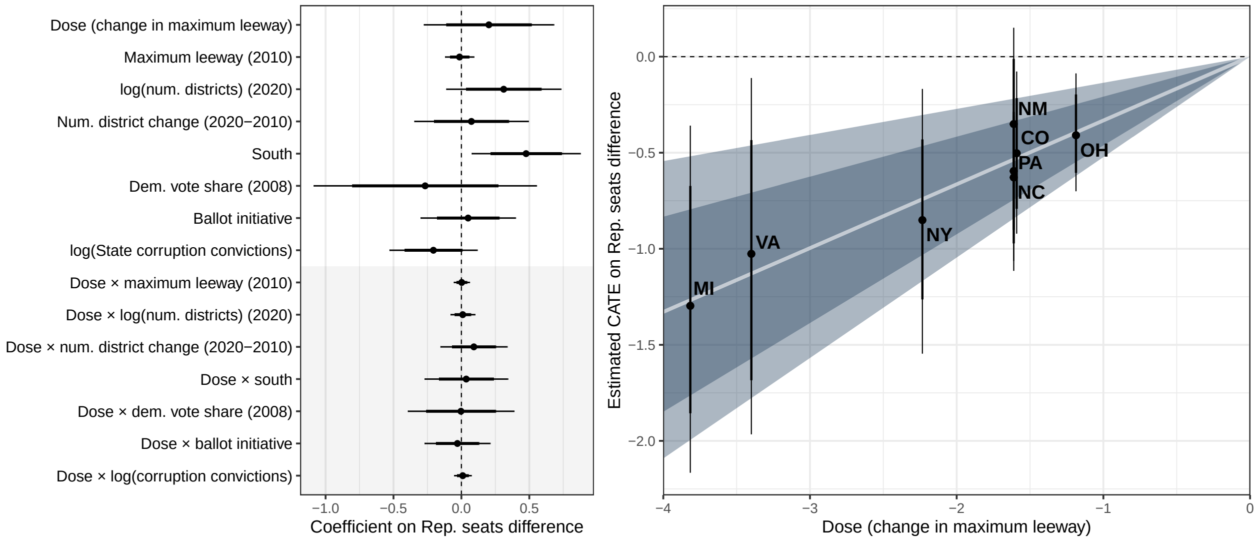

Figure 6: Fitted model coefficient estimates for the Republican seat outcome measure (left) using the maximum leeway treatment, and estimated conditional average treatment effects for each reformed state's covariate combination plotted against the state's dose (right). The model-based dose-response curve is underlaid in blue. 80% and 95% credible intervals are shown throughout.

coefficients of the estimation model discussed in Section 4 fit to the simulation-adjusted seats variable. All coefficients are shown on the scale of the outcome, where a positive coefficient indicates a relationship with a positive change in Republican seats from 2010 to 2020. The model's R^2 is around 0.4, highlighting the potential for the control variables to confound any observed correlation between reform and the outcome measure.

Given the fitted outcome model, we can estimate the conditional average treatment effect (CATE) for any given starting and final value of maximum leeway and any combination of covariates. The right panel of Figure 6 plots these CATEs for each state that experienced a change in maximum leeway between 2010 and 2020, using that state's specific covariates. There is a clear dose-response relationship between the decrease in maximum leeway and the decrease in the expected number of Republican seats.[12] A state with Michigan's dose (nearly 4) has an estimated CATE of around −1.25, with 95% credible interval (−2.1, −0.6), and so would be expected to gain just over 1 Democratic seat, on average. Similarly, a state with Ohio's dose (just over 1) and covariates would be expected to gain about half of a Democratic seat (0.2, 0.6). Motivated by this pattern, we also include in the right figure a dose-response curve calculated by calculating a CATE for dose level and averaging over the observed covariate distribution.[13]

Since the dose-response pattern is well-summarized by a linear relationship, we can summarize it by its slope, a quantity also known as the average causal response (ACR). The product of the ACR and a given dose correspond to the estimated causal effect of that dose, averaged across states.

Figure 7 presents the ACR for Republican seats and the other outcome variables, using both the maximum and realized leeway treatment values. The right side of the figure plots the ACR estimates in terms of the standard deviation of the respective outcome variable per unit change in treatment. This means that a point estimate of 0.5 is interpreted as an increase of 0.5 standard deviations of the outcome variable for a 1-unit change, or an increase of 2 standard deviations for a 4-unit change.

[12] This pattern was also observed in robustness checks that used a flexible Bayesian Additive Regression Trees (BART) model for estimation.

[13] Because of the interaction between the dose variable and the covariates (which do not have mean zero), the overall slope estimate and its variance differ from the estimate and variance of the coefficient on Dose in the linear model.

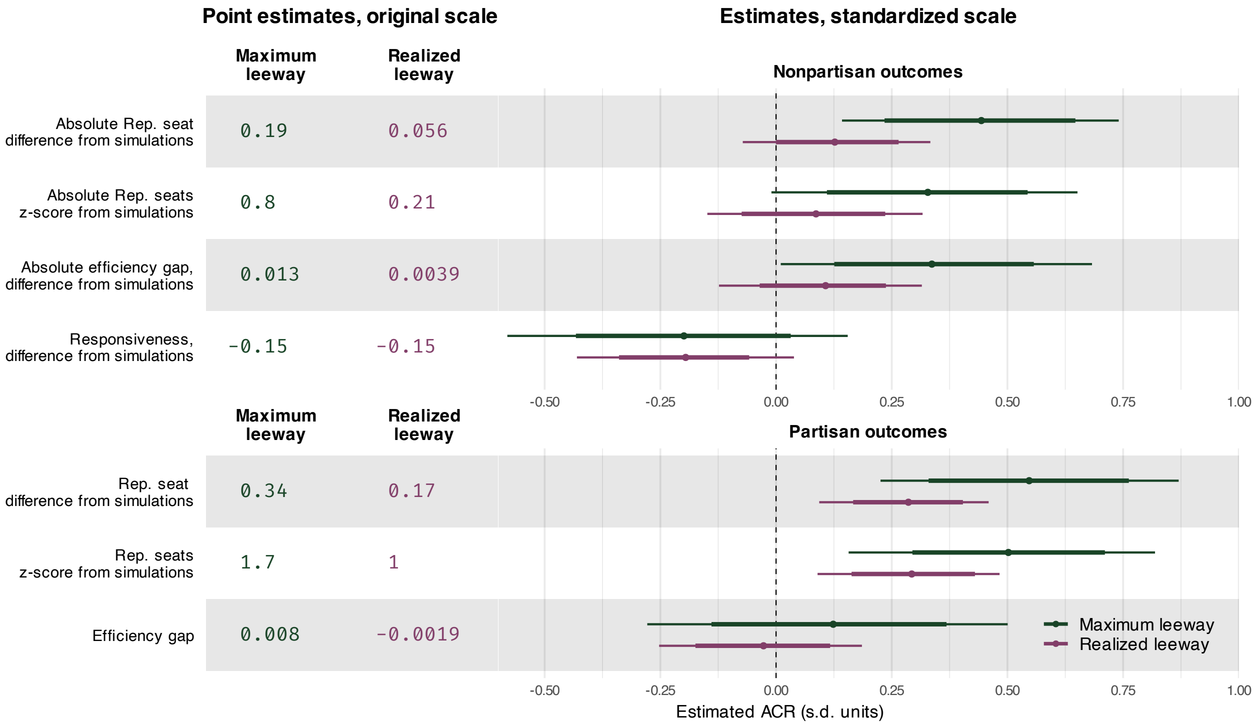

Figure 7: Average causal response (ACR) of leeway on redistricting outcomes. The points correspond to the mean estimated ACR, while the lines represent 80% and 95% credible intervals. Green corresponds to the maximum leeway treatment, while purple corresponds to the realized leeway treatment. The numbers in the columns display the mean ACR on each outcome's response scale. The estimates and intervals on the right are displayed in units of outcome standard deviations, to allow for comparability between outcomes. For partisan outcomes, a positive number indicates a pro-Republican effect and a negative number indicates a pro-Democratic effect for a positive dose.

Across the three nonpartisan outcome measures measuring the amount of gerrymandering, we find a consistent effect of leeway. The estimated ACRs are all positive, meaning an increase in leeway leads to an increased gap between the enacted plan and the simulated baseline; equivalently, states that reduce leeway through reform efforts are expected to rein in partisan gerrymandering. The magnitude of the effect is substantial: a reform effort like Michigan's, with a decrease in maximum leeway of 4 units, would be expected to reduce the difference in Republican seats from the baseline by about 0.8 seats (0.2, 1.3), or 3.2 simulation standard deviations (−0.1, 6.4)—enough to move an extreme partisan gerrymander to a fair plan. The ACRs for realized leeway are also positive, but are smaller in magnitude and have enough uncertainty that the 80% credible intervals cross zero. Thus there is some weak evidence that giving Democrats more control of the redistricting process by increasing their leeway or reducing Republican leeway would also reduce gerrymandering.

We also find that reforms increase electoral responsiveness. The point estimates of the ACRs for both maximum leeway and realized leeway agree; the posterior probability that the effect is negative is 86% and 96%, respectively. We estimate that a reduction in leeway of 4 units would increase responsiveness by 0.6 (−0.47, 1.75), meaning that a 1pp increase in a party's vote share would translate to an additional 0.6pp in seat share on top of the average 2pp increase in seat share across all states.

Another way to interpret the responsiveness effect is in terms of the share of competitive seats, where a seat that is counted as competitive in proportion to how close its baseline partisanship is to 50%. It turns out that this is equivalent to a linear rescaling of the responsiveness outcome.[14] In terms of competitive seats, we estimate that a reduction in leeway of 4 units would increase the share of competitive seats in an average state from 25% to 32% (19%, 46%).

In terms of partisan outcomes, the estimated ACRs on Republican seats are all positive, for both realized and maximum leeway. This is in line with an interpretation that *increasing* leeway benefits the party in control, as one might expect. Recall that the doses from Figure 4, for the realized leeway estimates, a 4 dose is approximately going from a Democratic legislature to an independent commission or an independent commission to a Republican legislature. For such a dose of 4, we would expect Republicans to gain around 0.69 seats (0.23, 1.12), or 6.9 simulation standard deviations (2.2, 11.3)—again, a substantial effect. We do not find a consistent effect on the signed efficiency gap outcome, which was not adjusted using the simulations.

6 Redistricting Commissions: Counterfactual Policy Analysis

What if redistricting reforms were adopted nationwide to constrain partisan actors? Redistricting commissions are the most commonly adopted map-drawing reform. A total of 15 states, including Michigan, New Jersey, and Arizona, currently use some form of commission. Advocates often argue that commissions can lead to fairer redistricting plans by further removing legislators from the process (e.g., Arts & Sciences Commission on the Practice of Democratic Citizenship, 2020). [15]

In this section, we conduct a counterfactual policy analysis about the nationwide adoption of redistricting commissions. Specifically, we use our fitted model and predict the electoral outcomes under a hypothetical scenario where every state adopts a particular type of redistricting commissions. Although each of the 15 states uses a redistricting commission with different structures and rules, our model is able to characterize these different commissions in terms of their partisan leeway, and use these treatment values to predict electoral outcomes under a given commission structure.

We study three types of commission structures currently enacted in several states: (1) a New York-style commission with a nonpartisan map drawer and several partisan veto points; (2) an Ohio-style reform with legislature-drawn map and several partisan and bipartisan veto points and stalemate procedures, including a bipartisan commission stalemate procedure; (3) and a Michigan-style reform, with a nonpartisan commission, no partisan veto points, and the potential for court review.

Figure 8 shows a seats-votes curve for each counterfactual commission scenario (the orange lines), which display the predicted number of total Democratic seats in the US House (y-axis) for a given national Democratic vote share (x-axis). We also include reference lines for actual and average simulated plans for both 2020 and 2010.

Seats-votes curves are often used to study a measure of partisan bias, operationalized as the difference in the share of seats and votes at a given point on the curve (Tufte, 1973; King and Browning, 1987; Katz et al., 2020). A seats-votes curve for which a 50% vote share translates exactly to a 50% seat share represents the baseline for partisan bais. The 2020 and 2010 simulated average lines represent estimates of the respective baselines for an unmanipulated seats-votes curve. These lines intersect the 50% vote line below the 50% seat line, since the underlying geographic distribution of Democratic votes nationally disadvantages the Democratic party.

We find that nationwide adoption of commissions would generally reduce partisan bias, but the extent of this reduction largely depends on details of commission structure. The model predicts that all commission structures

[14] Specifically, we measure the competitiveness of a single seat as the derivative of the election model's win probability for that seat at its baseline vote, divided by the maximum derivative, which is obtained at a 50/50 baseline vote. Based on our election model, a district with 50/50 baseline vote is counted as 1.0 competitive seats, and a district with 60/40 baseline vote is counted as just 0.14 competitive seats. This approach avoids arbitrary cutoffs in what counts as a competitive district.

[15] See also People Not Politicians Oregon or Ohio's Citizens Not Politicians

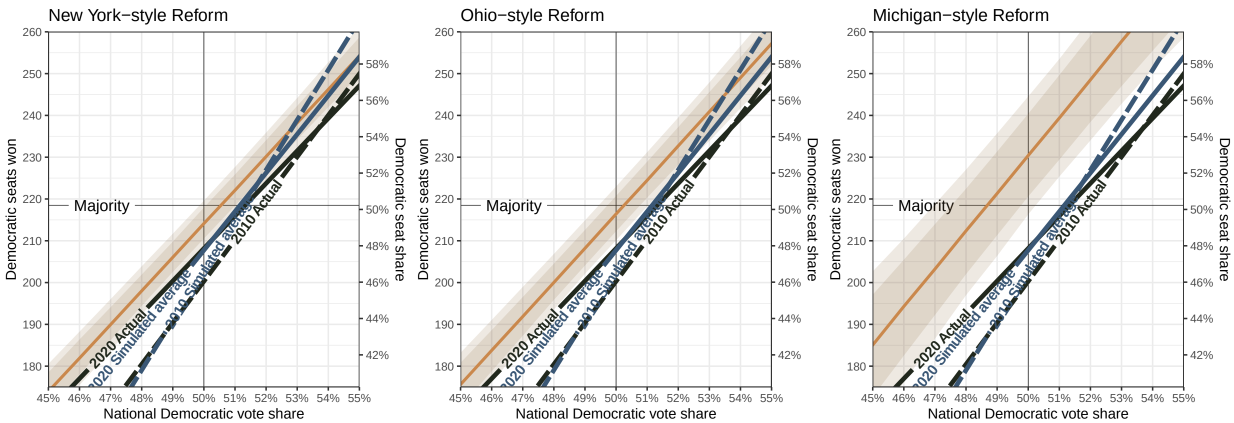

Figure 8: Commissions Can Reduce Partisan Bias. The figures show three predicted seats-votes curves if all US states adopted new redistricting institutions with: (1) a New York-style commission with a nonpartisan map drawer and several partisan veto points; (2) an Ohio-style legislature-drawn map and several partisan and bipartisan veto points; and (3) a Michigan-style reform, with a nonpartisan commission, no partisan veto points, and the potential for court review. Hypothetical commission structures are plotted as orange lines (with 80% and 95% credible intervals), with reference lines for actual and simulated plans for both 2020 and 2010. We find that commissions are predicted to reduce partisan bias at several relevant points like 50% and 51% along the seats-votes curve.

would reduce the existing Republican advantage in the House driven by both the inefficient geographic distribution of Democratic voters (Chen and Cottrell, 2016; Chen and Rodden, 2015) and net partisan manipulation favoring Republicans (Kenny et al., 2023).

We find that New York- and Ohio-style reforms produce relatively moderate improvements in partisan bias and responsiveness. The point estimates for the number of Democratic seats gained under these reforms is 5.9 for New York-style reforms, with 95% credible interval (0.28, 11.43), and 8.2 for Ohio-style reforms (1.9, 14.8). Responsiveness for New York- and Ohio-style reforms is also similar as for every 1pp increase in vote share, Democrats gain an estimated 8 seats with New York-style reforms, and 8.2 seats with Ohio-style reforms. These responsiveness values are only slightly higher than the actual responsiveness following 2020 redistricting. Ohio-style reforms appear more effective overall, though the substantive differences are small; we are more than 99% confident that Ohio-style reforms produce a larger effect on partisan bias than New York-style reforms and 85% confident that Ohio-style reforms increase responsiveness more than New York-style reforms do.

We find that Michigan-style reforms have the greatest effects in both responsiveness and Democratic seats, since partisan actors are the most constrained by the presence of a nonpartisan commission, no partisan veto points, and the potential for court review. We estimate that the number of Democratic seats gained under Michigan-style reforms is 22 with 95% credible interval (8.1, 36.4), and that for every 1pp increase in vote share, Democrats gain an estimated 9.1 seats (7, 11), an increase of 1.3 over the baseline. We are more than 99% confident that the Michigan-style reforms have a larger effect than the Ohio reforms, and more than 99% confident in a larger effect than the New York-style reforms. We are also 89% confident that the Michigan-style reform increases responsiveness.

Though all three proposed reforms reduce partisan bias, Michigan-style reforms suggest that much greater effects are produced when multiple nodes of the redistricting game constrain leeway by ensuring partisan nodes are nonpartisan, and do not precede partisan vetoes. Meanwhile, the examples of New York-style and Ohio-style reforms demonstrate how constraining partisan actors at different nodes, through different reforms, may produce substantively similar effects.

7 Conclusion

In this paper, we propose a methodology for estimating causal effects of complex institutional reforms and apply it to ascertain the impacts of redistricting reforms in the United States. Although redistricting reforms differ across states in important procedural details, we show how to obtain a theoretically informed univariate summary measure of such reforms through a game theoretic model. We then combine this approach with a standard causal inference research design strategy to draw a credible causal inference even with a limited sample size. We find that redistricting reforms reduce the partisan bias of enacted plans by constraining the leeway of partisan actors.

Our methodological approach enables us to go beyond the estimation of causal effects. Specifically, we conduct a counterfactual policy analysis to infer the consequences of different nationwide redistricting reforms. We find that instituting redistricting commissions generally reduces the current Republican advantage, but the details of commission structure matter. For example, Michigan-style reform would yield a much greater pro-Democratic effect than types of redistricting commissions adopted in Ohio and New York.

The key idea behind our methodology is also applicable to studies of other complex institutions. For example, most scholars analyze democratic institutions by using the Polity score, which is an 21-point scale index based on six "component variables" concerning executive recruitment, executive constraints, and political participation. Instead of simply summing these scores to obtain the overall Polity score, our proposed methodology suggests that researchers consider developing a more theoretically informed measure of democracy, based directly on their substantive questions of interest. We believe that such an approach preserves critical aspects of institutional complexity while improving the credibility of causal analysis.

References

Arora, A., Imai, K., Kenny, C. T., McCartan, C., Samarth, T., Simko, T., Wu, M., Wang, K., and Zhao, M. Y. (2024). Simulated redistricting plans for the 2010 redistricting cycle.

Arts & Sciences Commission on the Practice of Democratic Citizenship, A. A. o. (2020). Our common purpose.

Best, R. E., Lem, S. B., Magleby, D. B., and McDonald, M. D. (2021). Do Redistricting Commissions Avoid Partisan Gerrymanders? *American Politics Research*, page 1532673X211053216. Publisher: SAGE Publications Inc.

Cain, B. E. (2012). Redistricting Commissions: A Better Political Buffer? *the yale law journal*, page 37.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. (2024). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.

Canen, N. and Ramsay, K. (2023). Quantifying theory in politics: Identification, interpretation and the role of structural methods.

Canon, D. and Race, R. (1999). Representation: The unintended consequences of black majority districts.

Canon, D. T. (2022). Race and redistricting. *Annual Review of Political Science*, 25:509–528.

Carson, J. L. and Crespin, M. H. (2004). The effect of state redistricting methods on electoral competition in united states house of representatives races. *State Politics & Policy Quarterly*, 4(4):455–469.

Carson, J. L., Crespin, M. H., and Williamson, R. D. (2014). Reevaluating the effects of redistricting on electoral competition, 1972–2012. *State politics & policy quarterly*, 14(2):165–177.

Chen, J. and Cottrell, D. (2016). Evaluating partisan gains from Congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the U.S. House. *Electoral Studies*, 44:329–340.

Chen, J. and Rodden, J. (2015). Cutting Through the Thicket: Redistricting Simulations and the Detection of Partisan Gerrymanders. *Election Law Journal: Rules, Politics, and Policy*, 14(4):331–345. Publisher: Mary Ann Liebert, Inc., publishers.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, 1(1):451–488.

Douglas, J. A. (2014). The right to vote under state constitutions. *Vand. L. Rev.*, 67:89.

Ebanks, D., Katz, J. N., and King, G. (2023). How american politics ensures electoral accountability in congress. *Preprint*. Available at https://gking.harvard.edu/files/gking/files/acc.pdf.

Edwards, B., Crespin, M., Williamson, R. D., and Palmer, M. (2017). Institutional control of redistricting and the geography of representation. *The Journal of Politics*, 79(2):722–726.

Fraga, B. L., Moskowitz, D. J., and Schneer, B. (2022). Partisan alignment increases voter turnout: Evidence from redistricting. *Political behavior*, 44(4):1883–1910.

Gelman, A., Jakulin, A., Pittau, M. G., and Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4):1360–1383.

Gelman, A. and King, G. (1994). A unified method of evaluating electoral systems and redistricting plans. *American Journal of Political Science*, pages 514–554.

Golub, G. H. and Welsch, J. H. (1969). Calculation of gauss quadrature rules. *Mathematics of computation*, 23(106):221–230.

Gordon, S. C. and Yntiso, S. (0). Base rate neglect and the diagnosis of partisan gerrymanders. *Election Law Journal: Rules, Politics, and Policy*, 0(0):null.

Grofman, B. and Handley, L. (1991). Identifying and remedying racial gerrymandering. *JL & Pol.*, 8:345.

Katz, J. N., King, G., and Rosenblatt, E. (2020). Theoretical foundations and empirical evaluations of partisan fairness in district-based democracies. *American Political Science Review*, 114(1):164–178.

Keena, A., Latner, M., McGann, A. J. M., and Smith, C. A. (2021). *Gerrymandering the states: Partisanship, race, and the transformation of American federalism*. Cambridge University Press.

Kenny, C. T., McCartan, C., Simko, T., Kuriwaki, S., and Imai, K. (2023). Widespread partisan gerrymandering mostly cancels nationally, but reduces electoral competition. *Proceedings of the National Academy of Sciences*, 120(25):e2217322120.

King, G. and Browning, R. X. (1987). Democratic representation and partisan bias in congressional elections. *American Political Science Review*, 81(4):1251–1273.

Komisarchik, M. and White, A. (2021). Throwing away the umbrella: Minority voting after the supreme court's shelby decision. Technical report, Working Paper. https://arwhite. mit. edu/sites/default/files/images ....

Levitsky, S. and Ziblatt, D. (2023). *Tyranny of the minority: Why American democracy reached the breaking point*. Crown.

Little, A. T. and Meng, A. (2023). Measuring democratic backsliding. *PS: Political Science & Politics*, pages 1–13.

Lundberg, I., Johnson, R., and Stewart, B. M. (2021). What is your estimand? defining the target quantity connects statistical evidence to theory. *American Sociological Review*, 86(3):532–565.

Madison, J. (1788a). The apportionment of members among states. *The New York Packet*.

Madison, J. (1788b). The same subject continued: The House of Representatives. *The New York Packet*.

McCartan, C. and Imai, K. (2023). Sequential Monte Carlo for sampling balanced and compact redistricting plans. *Annals of Applied Statistics*, 17(4):3300–3323.

McCartan, C. and Kenny, C. T. (2022). Individual and differential harm in redistricting. *SocArXiv preprint*.

McCartan, C., Kenny, C. T., 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):689.

McCarty, N. (2019). *Polarization: What everyone needs to know*. Oxford University Press.

MIT Election Data and Science Lab (2017). U.S. House 1976–2020.

Nelson, M. (2023). Independent redistricting commissions are associated with more competitive elections. *PS: Political Science & Politics*, 56(2):207–212.

Parker, B. (2023). Polarization in state supreme courts, 1980-2020. *Available at SSRN 4430338*.

Polsby, D. D. and Popper, R. D. (1991). The third criterion: Compactness as a procedural safeguard against partisan gerrymandering. *Yale L. & Pol'y Rev.*, 9:301.

Rudensky, Y. (2023). Status of partisan gerrymandering litigation in state courts. *State Court Report*.

Stan Development Team (2024). Stan modeling language users guide and reference manual, version 2.32.

Stephanopoulos, N. O. and McGhee, E. M. (2015). Partisan gerrymandering and the efficiency gap. *U. Chi. L. Rev.*, 82:831.

Tufte, E. R. (1973). The relationship between seats and votes in two-party systems. *American Political Science Review*, 67(2):540–554.

U.S. Department of Justice Civil Rights Division (2013). Jurisdictions previously covered by section 5.

Wang, S. S.-H., Ober Jr, R. F., and Williams, B. (2019). Laboratories of democracy reform: State constitutions and partisan gerrymandering. *U. Pa. J. Const. L.*, 22:203.

Warshaw, C., McGhee, E., and Migurski, M. (2022). Districts for a New Decade—Partisan Outcomes and Racial Representation in the 2021–22 Redistricting Cycle. *Publius: The Journal of Federalism*, 52(3):428–451.

A Details of institutional coding

This appendix details the coding of each state's redistricting processes. The actual coded variables for each state in 2010 and 2020 are shown in Table A2.

A.1 Summary of encoding

Below, we provide a simplified summary of the coding, prior to the full details of the coding scheme.

A.2 Variable coding

Drawer. What kind of body draws the maps?

- **Legislature**: The state legislature directly draws the maps. *Example:* Florida 2020.
- **Commission**: Embodies any type of commission, including independent commission, bipartisan commission, or a legislative commission. *Example:* California 2020 had an independent commission.
- **N/A**: There is no redistricting. *Example:* Wyoming 2020 only has one congressional district and so does not redistrict.

Drawer Control. Who effectively controls the map-drawing body?

- **Democrats**. *Example:* Nevada 2020 had Democratic majorities in both chambers of the state legislature.
- **Republicans**. *Example:* North Carolina 2020 had Republican majorities in both chambers of the state legislature.
- **Split**: The map is drawn by the legislature, and partisan control of the legislature is split across both chambers. *Example:* Minnesota 2020 had a Democratic majority in the state house and a Republican majority in the state senate, so the legislature was under split partisan control.
- **Nonpartisans**: The map-drawing body is composed of people chosen through nonpartisan or explicitly bipartisan means. *Example:* California 2020 had an independent commission.
- **N/A**

Veto 1. Who has veto power over the proposed maps?

- **Legislature**: The legislature can choose to adopt or substitute their own maps for those presented to it by the initial map drawers. *Example:* Iowa 2020. The legislature can veto the commission's maps.
- **Governor**. *Example:* Alabama 2020.
- **Voters**: Maps can be subject to public referendum. *Example:* California 2020.
- **N/A**: No body has veto authority, or the map drawer has the votes/ability to override vetoes. *Example:* Kentucky 2020. Though the governor has statutory authority to veto maps from the legislature, the legislature had a Republican supermajority that overrode the Democratic governor's veto.

Veto 1 Control. Who controls the body with veto 2 power?

- **Democrats**. *Example:* Illinois 2020. The Democratic governor can veto the legislature's plan.
- **Republicans**. *Example:* Georgia 2020. The Republican governor can veto the legislature's plan.
- **Split**: The veto power is held by the legislature, which is split in its partisan control across both of its chambers. *Example:* Washington 2010. The legislature has veto power over the commission's maps, but the legislature was under split partisan control.

- **N/A**: No body has veto authority, or the map drawer has the votes/ability to override vetoes. *Example:* Tennessee 2020. The legislature has a Republican supermajority, and though the Republican governor has the statutory authority to veto maps, a supermajority renders this power effectively moot.

Veto 2. Who else has veto power over the proposed maps?

- **Governor**. *Example:* Iowa 2020. After the legislature receives maps from the legislature and approves a set of maps, the governor can veto the redistricting bill.
- **N/A**

Veto 2 Control. Who controls the body with veto 2 power?

- **Democrats**. *Example:* Maine 2020. The Democratic governor can veto the plan from the commission which was approved by the legislator.
- **Republicans**. *Example:* Iowa 2020. The Republican governor can veto the plan from the commission which was approved by the legislator.
- **N/A**: No body has veto 2 authority, or the map drawer has the votes/ability to override vetoes. *Example:* Utah 2020. The legislature (which has veto power over the commission) has a Republican supermajority, and though the Republican governor has the statutory authority to veto maps, a supermajority renders this power effectively moot.

Court Review. Is there a legal avenue for map challenges to partisan gerrymandering in state court? (Rudensky, 2023; Wang et al., 2019; Douglas, 2014)

- **Yes**: Clear enabling provisions in state law or constitution OR a legal challenge brought at any point in the past which was not dismissed. *Example:* New Mexico 2020. A state court accepted (but ultimately denied) a lawsuit from the New Mexico GOP challenging Congressional lines as a partisan gerrymander.
- **Maybe**: Possible enabling provisions in state law or constitution AND no legal challenge brought at any point. *Example:* South Carolina 2020.
- **No**: No enabling provisions in state law or constitution; or legal challenge clearly dismissed. *Example:* Kansas 2020. The State Supreme Court ruled that the use of partisan factors in redistricting is constitutionally permissible.
- **N/A**

Note that court review is a hidden variable that is not clear until a legal challenge is actually brought forth and either succeeds or is rejected. Indeed, even the specter of litigation may influence how other actors in redistricting approach the drawing of maps. When legal challenges were brought forth for the first time in a state in the 2020 redistricting cycle and the relevant laws under which the court ruled were in place in the 2010 redistricting cycle, we retroactively apply the permissibility of court review to the 2010 cycle as well under the assumption that had a challenge been brought in 2010, it would have proceeded the same way.

Court Control. Which party controls the state supreme court (or equivalent)? (Parker, 2023)

- **Democrats**. *Example:* Illinois 2020.
- **Republicans**. *Example:* Indiana 2020.

- **Split**: Equally divided court during the window for which challenges could have been brought for 2012 and 2022 maps. *Example:* Nevada 2020.
- **Nonpartisans**: Supreme court is truly nonpartisan, not just having nonpartisan elections. *Example:* South Carolina 2020.
- **N/A**

Stalemate 1. Who decides the outcome if the initial map drawer has a stalemate?

- **Court**: Court has statutory power to intervene in the event of a stalemate. *Example:* New Jersey 2020.
- **Commission**. *Example:* Oregon 2020.
- Commission Staff. *Example:* Colorado 2020.
- **Unclear**: no clear statutory or constitutional provision. *Example:* Louisiana 2020.
- **N/A**

Stalemate 1 Control. Who controls the initial stalemate outcome?

- **Democrats**. *Example:* Maine 2020. The Maine Supreme Court, which can intervene in stalemates, is controlled by Democrats.
- **Republicans**. *Example:* Indiana 2020. The Indiana Supreme Court, which can intervene in stalemates, is controlled by Republicans.
- **Split**: Stalemates are broken by a body that must have bipartisan support. *Example:* Ohio 2020. Stalemates are broken by a backup commission with the support of at least two members from each party (see exceptions below.)
- **Nonpartisans**: includes any type of commission selected in a nonpartisan or explicitly bipartisan way. *Example:* Oregon 2020.
- **N/A**

Stalemate 2. Who decides the outcome if the first stalemate-breaking body itself has a stalemate?

- **Court**: Court has statutory power to intervene in the event of a stalemate. *Example:* Connecticut 2020.
- **Legislature**. *Example:* Ohio 2020.
- **Unclear**. *Example:* Minnesota 2020.
- **N/A**

Stalemate 2 Control. Who controls the second stalemate outcome?

- **Democrats**. *Example:* Connecticut 2020.
- **Republicans**. *Example:* Ohio 2020.
- **N/A**

Final Drawer. Who drew the map in place for the 2012 or 2022 election cycle?

- **Legislature**: Includes re-drawing a remedial map after court challenges. *Example:* South Carolina 2020.
- **Commission**: Includes re-drawing a remedial map after court challenges. *Example:* Washington 2020.
- **Governor**. No example available.

- Court master: Court ordered a special master to draw the final plan. *Example:* Pennsylvania 2020.
- Court D remedy: Court picked a remedial map from a Democratic-affiliated plaintiff or intervenor. *Example:* Colorado 2020.
- Court R remedy: Court picked a remedial map from a Republican-affiliated plaintiff or intervenor. No example available.
- **N/A**

Preclearance. Was any part of the state subject to DOJ preclearance in 2010 (prior to Shelby County v. Holder)? (U.S. Department of Justice Civil Rights Division, 2013)

- **Yes**. *Example:* Georgia 2010.
- **No**. *Example:* Massachusetts 2010.

A.3 Notes on coding

Most criteria are straightforward, and the information for coding can easily be found in public data for each state. Furthermore, most states have similar redistricting policies that are easily classified under the categories we defined earlier. However, there are unique exceptions that we will treat below. In particular, we coded the below issues as follows.

Veto overrides.

- In Kansas 2020, Kentucky 2020, Louisiana 2020, Maryland 2010, Maryland 2020, Massachusetts 2010, Massachusetts 2020, Montana 2020, New York 2020, Rhode Island 2010, and Rhode Island 2020, the legislature had the necessary supermajority to override any vetoes, so we coded the governor's veto as N/A.

Unique commission compositions.

- In Idaho in 2010 and 2020, commissions are re-formed upon stalemate. Members are picked by both parties in a way that forms a 3-3 even split from both parties, except that the State Supreme Court picks the members of the commission if the parties don't decide in time. We code the control of this commission as having a partisan split.
- In Indiana in 2010 and 2020, stalemates are broken by a five-member backup commission composed of the majority leader from each house, the chair of the redistricting committee from each house, and a state legislator appointed by the governor. Since both houses and the governorship were controlled by Republicans, we coded this commission as being controlled by the Republican party.
- In Michigan 2020, if the commission does not adopt a final plan, a plan is randomly chosen. We code this as a stalemate 1 breaker by the nonpartisan commission.
- In Ohio 2020, if the legislature fails to redistrict by bipartisan majorities, the work falls to a commission that must approve a map by a bipartisan vote. The commission is composed of appointees by state executive officers and legislative parties. If the commission cannot do so, it can pass a map by a majority vote without bipartisan support but the map only lasts two election cycles instead. This is encoded as a draw by a split state legislature with the first stalemate broken by a nonpartisan commission and the second stalemate broken by Republicans, since Republicans control all state executive offices and have a functional majority on the commission.

Courts.

- In Kansas in 2010, the legislature failed to redistrict, and a federal court stepped in to draw the maps instead. But since this was not done under any statutory power and only done once, we consider this an exception to the normal redistricting process in Kansas rather than a procedurally institutionalized stalemate breaker. Similarly in Wisconsin in 2020, the State Supreme Court approved "least change" maps drawn by the governor after a stalemate, but we coded the stalemate as unclear for the same reasons.
- On the other hand, in Minnesota in 2010 and 2020 we have coded the court as a stalemate breaker because it has consistently intervened in previous stalemates.
- In Minnesota in 2020, a court-appointed panel drew the final map. We code this as equivalent to that of a court master.

Relevant history that did not affect encoding.

- In North Carolina, the 2012 map was ultimately struck down but only after the election. Furthermore, the 2022 map was an interim plan only used for the 2022 election.
- In Virginia in 2010, the state legislature initially deadlocked, but drew a map in time after state legislative elections in 2011. Thus the stalemate procedure for 2010 was coded as unclear.

Table A2: Institutional coding for all states.

State	Year	Drawer	Drawer control	Veto 1	Veto 1 ctrl.	Veto 2	Veto 2 ctrl.	Court review?	Court control	Stalemate 1	Stalemate 1 ctrl.	Stalemate 2	Stalemate 2 ctrl.	Preclearance
Alabama	2010	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Alabama	2020	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Arizona	2010	commission	nonpartisans	NA	NA	NA	NA	maybe	republicans	unclear	NA	NA	NA	yes
Arizona	2020	commission	nonpartisans	NA	NA	NA	NA	maybe	republicans	unclear	NA	NA	NA	yes
Arkansas	2010	legislature	democrats	governor	democrats	NA	NA	maybe	democrats	unclear	NA	NA	NA	yes
Arkansas	2020	legislature	republicans	governor	republicans	NA	NA	maybe	republicans	unclear	NA	NA	NA	yes
California	2010	commission	nonpartisans	voters	NA	NA	NA	yes	republicans	court	democrats	unclear	NA	yes
California	2020	commission	nonpartisans	voters	NA	NA	NA	yes	democrats	court	republicans	unclear	NA	yes
Colorado	2010	legislature	split	governor	democrats	NA	NA	yes	democrats	court	democrats	unclear	NA	no
Colorado	2020	commission	nonpartisans	court	democrats	NA	NA	yes	democrats	commission staff	nonpartisans	unclear	NA	no
Connecticut	2010	legislature	split	NA	NA	NA	NA	no	democrats	commission	nonpartisans	court	democrats	no
Connecticut	2020	legislature	split	NA	NA	NA	NA	no	democrats	commission	nonpartisans	court	democrats	no
Florida	2010	legislature	republicans	governor	republicans	NA	NA	yes	democrats	unclear	NA	NA	NA	yes
Florida	2020	legislature	republicans	governor	republicans	NA	NA	yes	republicans	unclear	NA	NA	NA	yes
Georgia	2010	legislature	republicans	governor	republicans	NA	NA	no	democrats	unclear	NA	NA	NA	yes
Georgia	2020	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Hawaii	2010	commission	nonpartisans	NA	NA	NA	NA	yes	democrats	unclear	NA	NA	NA	no
Hawaii	2020	commission	nonpartisans	NA	NA	NA	NA	yes	democrats	unclear	NA	NA	NA	no
Idaho	2010	commission	split	NA	NA	NA	NA	yes	republicans	commission	split	unclear	NA	no
Idaho	2020	commission	split	NA	NA	NA	NA	yes	republicans	commission	split	unclear	NA	no
Illinois	2010	legislature	democrats	governor	democrats	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Illinois	2020	legislature	democrats	governor	democrats	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Indiana	2010	legislature	republicans	governor	republicans	NA	NA	maybe	republicans	commission	republicans	unclear	NA	no
Indiana	2020	legislature	republicans	governor	republicans	NA	NA	maybe	republicans	commission	republicans	unclear	NA	no
Iowa	2010	commission	nonpartisans	legislature	republicans	governor	republicans	no	republicans	legislature	republicans	unclear	NA	no
Iowa	2020	commission	nonpartisans	legislature	republicans	governor	republicans	no	republicans	legislature	republicans	unclear	NA	no
Kansas	2010	legislature	republicans	governor	republicans	NA	NA	no	nonpartisans	unclear	NA	NA	NA	no
Kansas	2020	legislature	republicans	governor	NA	NA	NA	no	nonpartisans	unclear	NA	NA	NA	no
Kentucky	2010	legislature	split	governor	democrats	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Kentucky	2020	legislature	republicans	governor	NA	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Louisiana	2010	legislature	republicans	governor	republicans	NA	NA	no	democrats	unclear	NA	NA	NA	yes
Louisiana	2020	legislature	republicans	governor	NA	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Maine	2010	commission	nonpartisans	legislature	split	governor	republicans	no	democrats	court	democrats	NA	NA	no
Maine	2020	commission	nonpartisans	legislature	split	governor	democrats	no	democrats	court	democrats	NA	NA	no
Maryland	2010	legislature	democrats	governor	NA	voters	NA	yes	democrats	unclear	NA	NA	NA	yes
Maryland	2020	legislature	democrats	governor	NA	voters	NA	yes	nonpartisans	unclear	NA	NA	NA	yes
Massachusetts	2010	legislature	democrats	governor	NA	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Massachusetts	2020	legislature	democrats	governor	NA	NA	NA	maybe	nonpartisans	unclear	NA	NA	NA	no
Michigan	2010	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Michigan	2020	commission	nonpartisans	NA	NA	NA	NA	yes	democrats	commission	nonpartisans	NA	NA	yes
Minnesota	2010	legislature	republicans	governor	democrats	NA	NA	maybe	republicans	court	republicans	unclear	NA	no
Minnesota	2020	legislature	split	governor	democrats	NA	NA	maybe	democrats	court	democrats	unclear	NA	no
Mississippi	2010	legislature	split	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Mississippi	2020	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Missouri	2010	legislature	republicans	governor	NA	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Missouri	2020	legislature	republicans	governor	republicans	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Montana	2020	commission	nonpartisans	NA	NA	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Nebraska	2010	legislature	republicans	governor	republicans	NA	NA	no	democrats	unclear	NA	NA	NA	no
Nebraska	2020	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	no
Nevada	2010	legislature	democrats	governor	republicans	NA	NA	no	nonpartisans	unclear	NA	NA	NA	no
Nevada	2020	legislature	democrats	governor	democrats	NA	NA	no	split	unclear	NA	NA	NA	no
New Hampshire	2010	legislature	republicans	governor	democrats	NA	NA	no	democrats	unclear	NA	NA	NA	no
New Hampshire	2020	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	no
New Jersey	2010	commission	nonpartisans	NA	NA	NA	NA	no	republicans	court	republicans	NA	NA	no
New Jersey	2020	commission	nonpartisans	NA	NA	NA	NA	no	democrats	court	democrats	NA	NA	no
New Mexico	2010	legislature	democrats	governor	republicans	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
New Mexico	2020	legislature	democrats	governor	democrats	NA	NA	yes	democrats	unclear	NA	NA	NA	no
New York	2010	legislature	split	governor	democrats	NA	NA	no	republicans	unclear	NA	NA	NA	yes
New York	2020	commission	nonpartisans	legislature	democrats	governor	NA	yes	democrats	legislature	democrats	unclear	NA	yes
North Carolina	2010	legislature	republicans	NA	NA	NA	NA	maybe	republicans	unclear	NA	NA	NA	yes
North Carolina	2020	legislature	republicans	NA	NA	NA	NA	yes	democrats	unclear	NA	NA	NA	yes
Ohio	2010	legislature	republicans	governor	republicans	voters	NA	maybe	republicans	unclear	NA	NA	NA	no
Ohio	2020	legislature	split	governor	republicans	NA	NA	yes	republicans	commission	split	legislature	republicans	no
Oklahoma	2010	legislature	republicans	governor	NA	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Oklahoma	2020	legislature	republicans	governor	republicans	NA	NA	maybe	republicans	unclear	NA	NA	NA	no
Oregon	2010	legislature	split	governor	democrats	NA	NA	yes	democrats	unclear	NA	NA	NA	no
Oregon	2020	legislature	democrats	governor	democrats	NA	NA	yes	democrats	commission	nonpartisans	NA	NA	no
Pennsylvania	2010	legislature	republicans	governor	republicans	NA	NA	maybe	republicans	unclear	NA	NA	NA	no
Pennsylvania	2020	legislature	republicans	governor	democrats	NA	NA	yes	democrats	unclear	NA	NA	NA	no
Rhode Island	2010	commission	nonpartisans	legislature	democrats	governor	NA	no	split	unclear	NA	NA	NA	no
Rhode Island	2020	commission	nonpartisans	legislature	democrats	governor	NA	no	split	unclear	NA	NA	NA	no
South Carolina	2010	legislature	republicans	governor	republicans	NA	NA	maybe	nonpartisans	unclear	NA	NA	NA	yes
South Carolina	2020	legislature	republicans	governor	republicans	NA	NA	maybe	nonpartisans	unclear	NA	NA	NA	yes
Tennessee	2010	legislature	republicans	governor	NA	NA	NA	maybe	democrats	unclear	NA	NA	NA	no
Tennessee	2020	legislature	republicans	governor	NA	NA	NA	maybe	republicans	unclear	NA	NA	NA	no
Texas	2010	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Texas	2020	legislature	republicans	governor	republicans	NA	NA	no	republicans	unclear	NA	NA	NA	yes
Utah	2010	legislature	republicans	governor	NA	NA	NA	no	nonpartisans	unclear	NA	NA	NA	no
Utah	2020	commission	nonpartisans	legislature	republicans	governor	NA	no	nonpartisans	legislature	republicans	NA	NA	no
Virginia	2010	legislature	republicans	governor	republicans	NA	NA	no	nonpartisans	unclear	NA	NA	NA	yes
Virginia	2020	commission	nonpartisans	legislature	split	NA	NA	no	nonpartisans	court	nonpartisans	NA	NA	yes
Washington	2010	commission	nonpartisans	legislature	split	NA	NA	yes	nonpartisans	court	democrats	NA	NA	no
Washington	2020	commission	nonpartisans	legislature	democrats	NA	NA	yes	democrats	court	democrats	NA	NA	no
West Virginia	2010	legislature	democrats	governor	NA	NA	NA	no	democrats	unclear	NA	NA	NA	no
West Virginia	2020	legislature	republicans	governor	NA	NA	NA	no	republicans	unclear	NA	NA	NA	no
Wisconsin	2010	legislature	republicans	governor	republicans	NA	NA	no	nonpartisans	unclear	NA	NA	NA	no
Wisconsin	2020	legislature	republicans	governor	democrats	NA	NA	no	republicans	unclear	NA	NA	NA	no

B Redistricting Game Specification

The game tree is specified in Figure 2 and is described in Section 3.2. Here we detail the parametric specification of the moves by nature, starting with the post-enactment court review process.

B.1 Post-enactment processes

We decompose the court challenge process into five components, as follows:

1. The probability a legal challenge is possible, which depends on the data coding, is specified as

$$\texttt{pr\_chal\_poss} = \begin{cases} \texttt{chal\_poss\_conf} & \text{if Court Review is coded as "Yes"} \\ \texttt{chal\_poss\_maybe} & \text{if Court Review is coded as "Maybe"} \\ 1 - \texttt{chal\_poss\_conf} & \text{if Court Review is coded as "No"} \end{cases}$$

The prior on `chal_poss_conf` is Beta$(19, 1)$ and the prior on `chal_poss_maybe` is Beta$(6, 14)$. The effect of this setup is to follow the data closely but allow for some uncertainty in the possibility of court review, rather than, e.g., precluding any possibility of court challenge in a state because it was coded as "No." State courts may change their interpretations of state laws and constitutions in any given case.

2. The probability a challenge is made when one is possible, which depends on the extremity of the partisan bias of the first enacted plan, x. Let $F(x) = \pi^{-1} \arctan(x) + \frac{1}{2}$ be the CDF of a standard Cauchy distribution. The specification of this probability is then

$$\begin{aligned} \texttt{pr\_chal\_if\_poss}(x) &= F(a + bx^2) \\ a &= F^{-1}(\texttt{chal\_prob\_bias}_0) \\ b &= (F^{-1}(\texttt{chal\_prob\_bias}_2) - F^{-1}(\texttt{chal\_prob\_bias}_0))/2^2. \end{aligned}$$

The general form of $F(a + bx^2)$ means that the probability of a challenge will be U-shaped, asymptotically approaching 1 as the absolute partisan bias increases. For interpretability, the scale and shift coefficients are parametrized in terms of the probability of a challenge when the first enacted plan has bias 0 and the probability of a challenge when the first enacted plan has bias 2. These parameters receive Beta$(4, 16)$ and Beta$(17, 3)$ priors, respectively. 100 draws of the induced prior on `pr_chal_if_poss` are shown in Figure B1.

3. The probability a court sides with plaintiffs challenging the first enacted plan, i.e., that the court intervenes, which depends on the partisan bias of the first enacted plan as well as the partisan makeup of the state supreme court. We assume that courts with a partisan majority are more likely to intervene and find in favor of challengers when the partisan bias of the first enacted plan is opposite in sign to the court's partisan lean. Let $g(x; a) = x^2(0.7 + ac \cdot x + (0.2x)^2)$, with $c = \sqrt{4(2 \cdot 0.7)(12 \cdot 0.2^2)}/6$, which is a convex quartic with asymmetry controlled by a parameter a; the value c is set so that the discriminant of the second derivative is positive, i.e., that the quartic is indeed convex. Then the intervention probability is

$$\texttt{pr\_intervene}(x) = \begin{cases} bc \cdot g(-x; \texttt{interv\_asym}) + ac & \text{if Court Control is coded as "Democrats"} \\ bc \cdot g(-x; \texttt{interv\_asym}) + ac & \text{if Court Control is coded as "Republicans"} \\ ac + bcx^2 & \text{otherwise} \end{cases},$$

with

$$a = F^{-1}(\texttt{interv\_prob\_bias}_0)$$

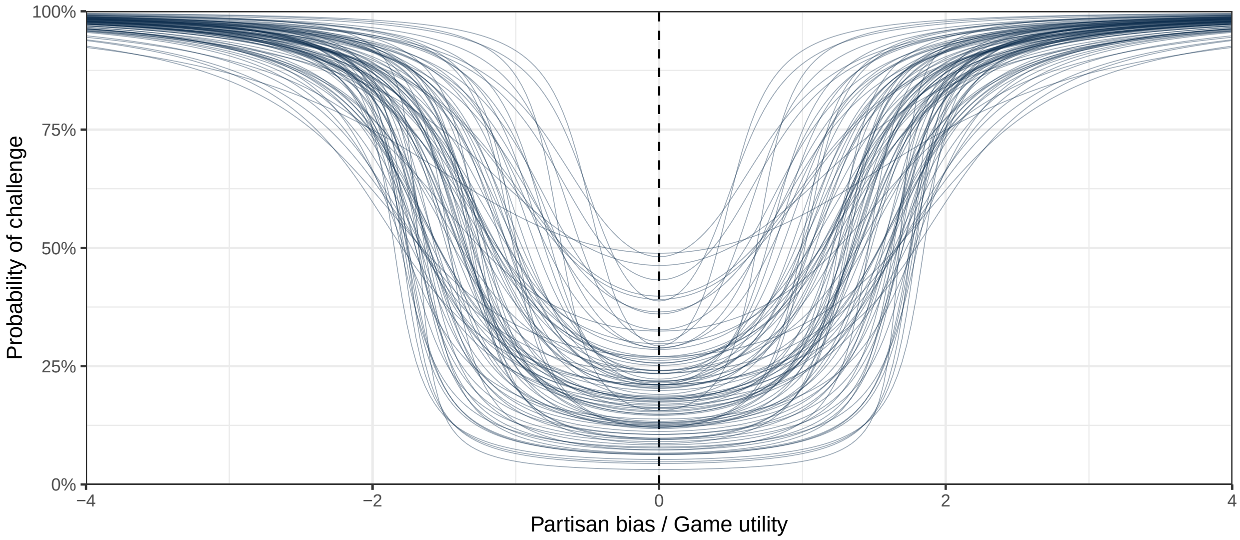

Figure B1: 100 draws from the prior on the `pr_chal_if_poss`(x) curve.

$$b = (F^{-1}(\texttt{interv\_prob\_bias}_2) - F^{-1}(\texttt{interv\_prob\_bias}_0))/2^2$$
$$c = \texttt{interv\_prob\_max}$$

The coefficients on the quartic g were adjusted so that the average value of g across the interval $[-4, 4]$ was close to the average value of x^2 across the same interval. Like the challenge probability, the general form of $F(ac + bcx^2)$ means that the probability of intervention will be U-shaped, asymptotically approaching $c =$ `inverv_prob_max` as the absolute partisan bias increases. The use of g means that the probability of intervention will be asymmetric around 0 for partisan courts. The maximum `inverv_prob_max` has a Beta$(18, 2)$ prior; the asymmetry parameter has a Beta$(4, 1.5)$ prior; the relative probability of intervention for a neutral plan, $\texttt{interv\_prob\_bias}_0$, receives a Beta$(4, 16)$ prior; and the relative probability of intervention for a plan with bias 2, $\texttt{interv\_prob\_bias}_2$, receives a Beta$(18, 2)$ prior. 100 draws of the induced prior on `pr_intervene` for each type of court control are shown in Figure B2.

4. The expected remedy a court orders when siding with plaintiffs, which depends on the partisan bias of the first enacted plan as well as the partisan makeup of the state supreme court, and is specified as

$$\texttt{court\_outcome}(x) = \frac{\texttt{out\_nonp\_bias}_2}{\arctan(2/2)} \cdot \arctan(x/2) + a \cdot \texttt{out\_nonp\_part\_adv}$$
$$a = \begin{cases} -1 & \text{if Court Control is coded as “Democrats”} \\ 1 & \text{if Court Control is coded as “Republicans”} \\ 0 & \text{otherwise} \end{cases}.$$

The `out_nonp_bias` parameter governs the partisan bias of a nonpartisan court remedy when the partisan bias of the first enacted plan is 2; it receives a folded Normal prior with mean 0 and standard deviation $\frac{1}{2}$, i.e., we simulate from $\mathcal{N}\left(0, \frac{1}{2}\right)$ and then take absolute values. The `out_nonp_part_adv` parameter controls the additional partisan lean of the remedy towards the party that controls the court, where applicable; it receives a folded Normal prior with mean 0 and standard deviation 0.4. Figure B3 shows 100 draws from the induced prior on `court_outcome` for each type of court control.

5. The probability and expected effect of a challenge based on the federal Voting Rights Act, which depends on whether a state was subject to DOJ preclearance pre-*Shelby*, as well as the partisan bias of the first

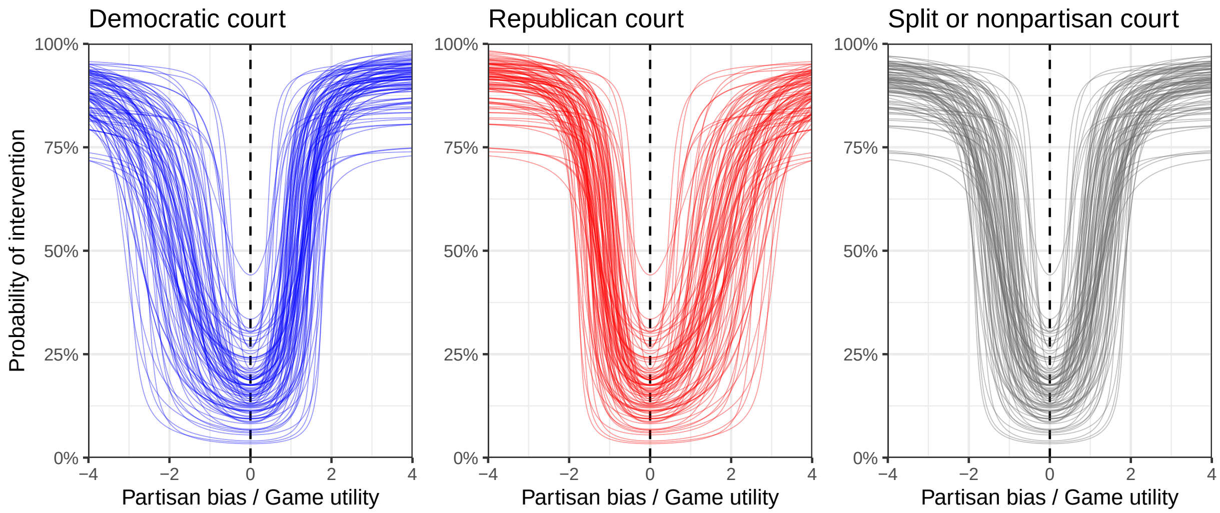

Figure B2: 100 draws from the prior on the `pr_intervene`(x) curve for each type of partisan court control.

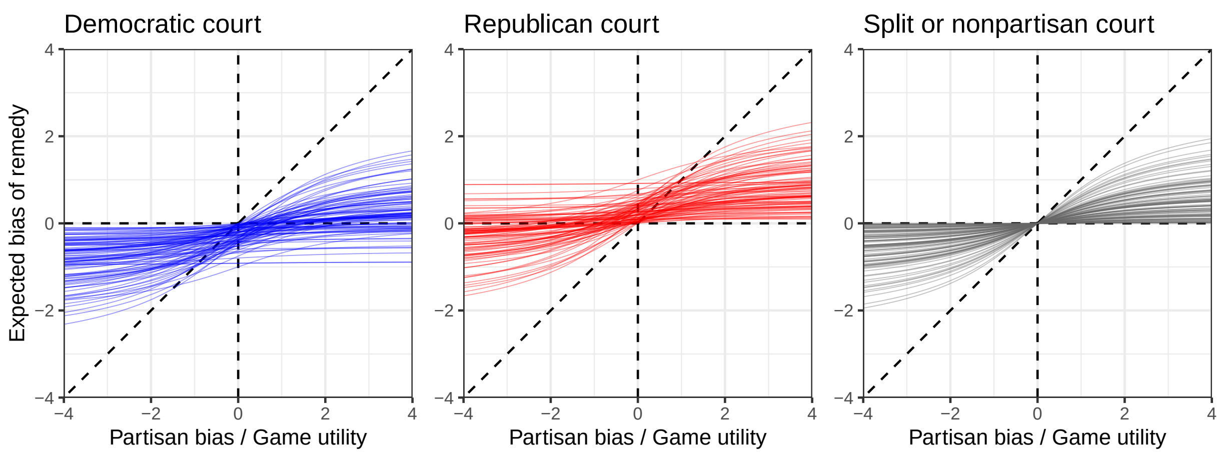

Figure B3: 100 draws from the prior on the `court_outcome`(x) curve for each type of partisan court control.

enacted plan. The probability of a VRA challenge is zero for non-preclearance states; for preclearance states it is

$$\begin{aligned} \texttt{pr\_vra\_chal}(x) &= F(a+bx) \\ a &= F^{-1}(\texttt{vra\_chal\_prob\_bias}_0) \\ b &= (F^{-1}(\texttt{vra\_chal\_prob\_bias}_2) - F^{-1}(\texttt{chal\_prob\_bias}_0))/2. \end{aligned}$$

The general form of $F(a+bx)$ means the challenge probability is asymmetrical: challenges are more likely for plans that favor Republicans. This follows from the high levels of racially polarized voting in the U.S. and specifically the strong preference for Democratic candidates by minority groups. The $\texttt{vra\_chal\_prob\_bias}_0$ parameter is the probability of a VRA challenge against a plan with no partisan bias; its prior is Beta$(2, 18)$. Similarly, $\texttt{vra\_chal\_prob\_bias}_2$, the challenge probability for a Republican-favoring plan, has a $\text{Beta}(9, 1)$ prior. The left panel of Figure B4 shows 100 draws from the induced prior on `pr_vra_chal`.

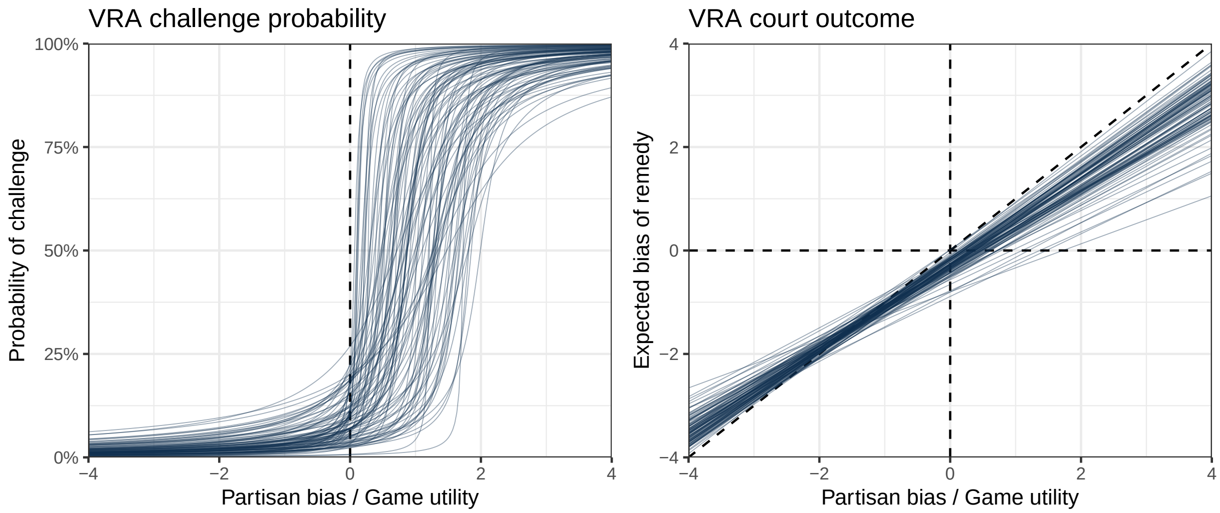

Figure B4: 100 draws from the prior on the $\texttt{pr\_vra\_chal}(x)$ curve (left) and the $\texttt{court\_vra\_outcome}(x)$ curve (right).

We assume the probability of a successful VRA claim does not depend on the partisan bias of the first enacted plan; this parameter is `vra_interv_prob` and receives a prior of $\text{Beta}(4, 1.5)$. The court-ordered remedy for a successful VRA challenge is specified as

$$\texttt{court\_vra\_outcome}(x) = \texttt{vra\_out\_slope}(x - \texttt{vra\_out\_breakeven}) + \texttt{vra\_out\_breakeven};$$

the intercept `vra_out_breakeven` has a $\mathcal{N}(-1.5, 0.5)$ prior and the slope `vra_out_slope` has a $\text{Beta}(16, 4)$ prior. So in expectation VRA remedies make Republican-favoring plans slightly more Democratic. The right panel of Figure B4 shows 100 draws from the induced prior on `court_vra_outcome`.

Putting these five pieces together, we can write the expected outcome of the post-enactment processes as

$$\begin{aligned}
\texttt{exp_court}(x) &= \texttt{pr_int_net}(x) \cdot \texttt{court_outcome}(x) \\
&\quad + (1 - \texttt{pr_int_net}(x)) \cdot \texttt{pr_vra_net}(x) \cdot \texttt{court_vra_outcome}(x) \\
&\quad + (1 - \texttt{pr_int_net}(x)) \cdot (1 - \texttt{pr_vra_net}(x)) \cdot x \\
\texttt{pr_int_net}(x) &= \texttt{pr_chal_poss} \cdot \texttt{pr_chal_if_poss}(x) \cdot \texttt{pr_intervene}(x) \\
\texttt{pr_vra_net}(x) &= \texttt{pr_vra_chal}(x) \cdot \texttt{vra_interv_prob}
\end{aligned}$$

Figure B5 shows 50 draws from the induced prior on $\texttt{exp\_court}(x)$ for each combination of court control, court review, and DOJ preclearance. We can see that the most important factor in shaping post-enactment outcomes is the presence of court review, as might be expected. When court review is uncertain or not present, court processes are expected to only moderately constrain outcomes.

B.2 Stalemate Process

When there is no stalemate process specified, when stalemate processes are exhausted, or when the stalemate process is controlled by neither party, the stalemate outcome is assumed to be a linear rescaling of the bias of

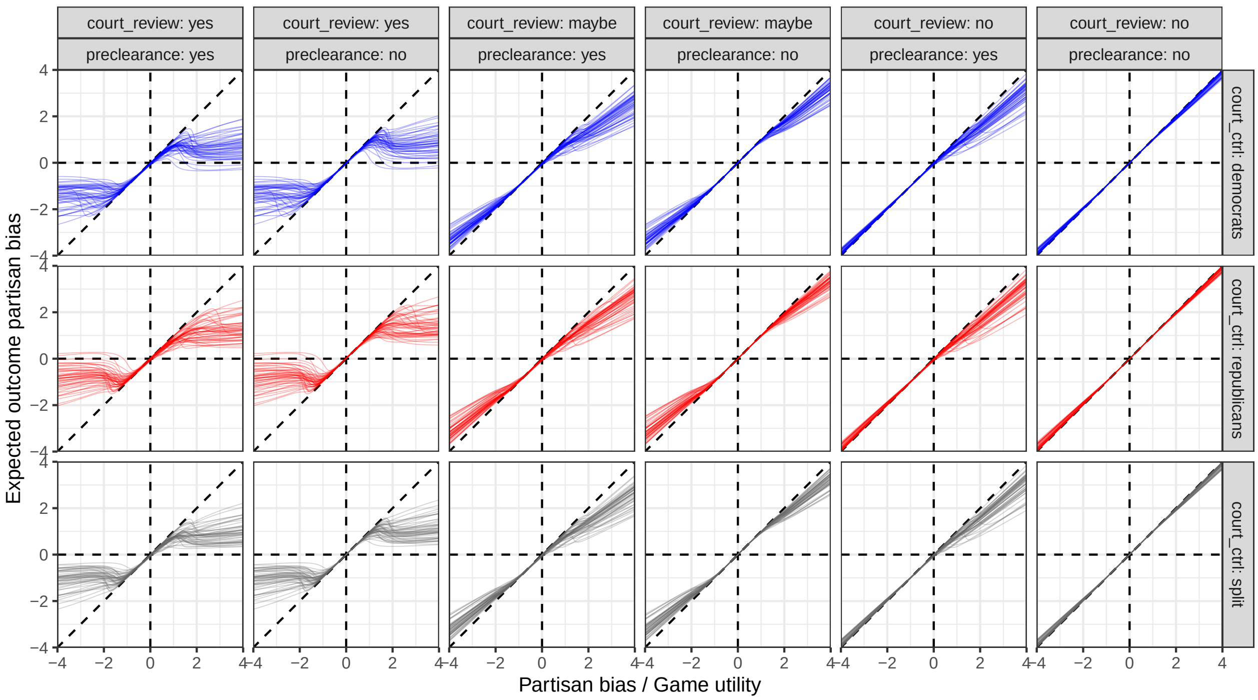

Figure B5: 50 draws from the prior on $\texttt{exp\_court}(x)$ for each combination of relevant procedural variables.

the plan drawn at the previous stage, x, which acts as a sort of anchoring point, and the partisan control of the initial map drawer. The specification is

$$\texttt{stalemate\_default}(x) = \begin{cases} \texttt{stale\_slope} \cdot x - \texttt{out\_nonp\_part\_adv} & \text{if the initial drawer is Democratic} \\ \texttt{stale\_slope} \cdot x + \texttt{out\_nonp\_part\_adv} & \text{if the initial drawer is Republican} \\ \texttt{stale\_slope} \cdot x & \text{otherwise} \end{cases},$$

with out_nonp_part_adv defined as above, and stale_slope receiving a $\text{Beta}(3, 17)$ prior. This reflects a status quo bias towards the last-drawn plan.

When stalemates are specified as being resolved by state courts, and the state court is partisan, the stalemate outcome is

$$\begin{cases} \texttt{stale\_slope} \cdot x - \texttt{out\_nonp\_part\_adv} & \text{if Court Control is coded as “Democrats”} \\ \texttt{stale\_slope} \cdot x + \texttt{out\_nonp\_part\_adv} & \text{if Court Control is coded as “Republicans”} \end{cases}$$

given an input partisan bias x.

B.3 Nonpartisan veto points

When veto control is coded as as NA (other than in the case of a governor with overridable veto) or a court has a veto, the veto is exercised with probability

$$\texttt{pr\_veto\_chal}(x) = \texttt{veto\_nonp\_prob\_max} \cdot \texttt{pr\_chal\_if\_poss}(x),$$

with the parameter veto_nonp_prob_max receiving a $\text{Beta}(3, 7)$ prior. So these veto players are unlikely to intervene, but relatively more likely if the proposed plan is more extreme

Split-control veto players are assumed to never exercise their veto, as are governors that can be overridden.

B.4 Split-control and nonpartisan map drawers

Map drawers with split partisan control are assumed to stalemate with probability controlled by parameter `stale_split_prob`, which has a $\text{Beta}(3, 5)$ prior. When they don't stalemate, split-control map drawers are assumed to propose a map with the same partisan bias as one drawn by a nonpartisan map drawer.

In Round 1, nonpartisan map drawers are assumed to propose a plan with partisan bias of 0. In Round 2—after one veto has already occurred—nonpartisan map drawers are assumed to shift the bias of their proposed plan in the direction of the average partisanship of the veto players. Specifically, if x is the proposal from the previous stage, then the proposal in Round 2 is

$$\texttt{exp\_drawer\_r2\_nonp}(x) = x + \texttt{veto\_party} \cdot \texttt{veto\_nonp\_shift}$$

$$\texttt{veto\_party} = \begin{cases} -1 & \text{if both veto players are Democratic} \\ 1 & \text{if both veto players are Republican} \\ -\frac{1}{2} & \text{if one veto player is Democratic and the other is nonpartisan or absent} \\ \frac{1}{2} & \text{if one veto player is Republican and the other is nonpartisan or absent} \\ 0 & \text{otherwise} \end{cases}.$$

The `veto_nonp_shift` parameter controls the amount of the shift and receives a $\mathcal{N}(0.65, 0.3)$ prior.

C Additional Outcomes

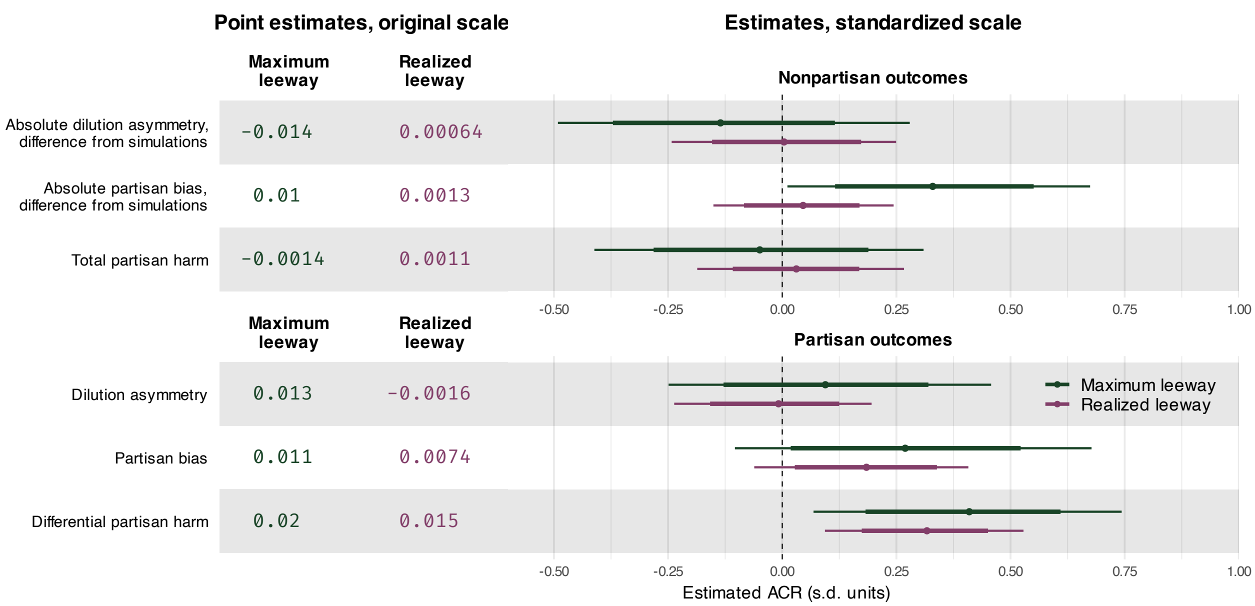

Figure C6: Average causal response of one dose of "leeway" on redistricting outcomes. The points correspond to the mean estimated ACR, while the lines represent 95% confidence intervals. Green corresponds to the party-blind (worst case) treatment, while purple corresponds to the party-signed treatment. The numbers in the columns display the mean ACR on each outcome's response scale. The points and lines are displayed in standard deviation units to allow for comparability between outcomes. For partisan outcomes, a positive number indicates a pro-Republican effect and a negative number indicates a pro-Democratic effect for a positive dose.

For completeness, we consider a series of alternative outcomes which measure the partisan bias of redistricting plans. We first consider an alternative to the efficiency gap, the dilution asymmetry (Gordon and Yntiso, 0). This measures the difference in the percentage of each party's votes that are wasted (cf. the number of wasted votes used in the calculation of the efficiency gap). Similar to the efficiency gap, we consider both the absolute value of differences from simulations and signed raw values for the dilution asymmetry.

A series of partisan bias measures are based on the seats-votes curve. Under a fair plan, the seats-votes curve should be symmetric (Katz et al., 2020). We consider two groups of measures based on the seats-votes curve. First, we consider the "partisan bias" which measures the deviation from symmetry in the seats-votes curve. We use both the raw signed value and absolute value of the difference from simulations of the partisan bias.

Finally, we consider total and differential individual harm, which measures how a redistricting plan negatively affects the ability for a Democratic or Republican voter to elect a candidate of their choice, relative to simulated redistricting plans (McCartan and Kenny, 2022). Differential partisan harm is a signed measure where positive scores indicate Democrats are harmed at a higher rate than Republican voters. The total partisan harm of a plan is a nonpartisan measure, which represents the total harm to both parties.

D Election modeling

The following information about election modeling has been adapted from the methods section and supplementary materials of Kenny et al. (2023) and is included here for completeness.

An election model uses observed past election results to quantify the uncertainty over future election results which take place under a different set of redistricting plans.

Our goal is to estimate the partisanship of each district in the enacted plan and also for alternative districts, with different geographic configurations, in each of the 5,000 simulated plans. To do this, we use precinct-level election data. We estimate the precinct-level baseline partisanship for 2010 using 2008 presidential election data and the baseline for 2020 as an average of the 2016 and 2020 presidential elections. We use previous presidential elections because the same candidate is on the ballot across the entire nation, unlike for Senate or House races. This practice is also adopted by many elections analysts and is used in the Cook Partisan Voting Index.

Specifically, we take the mean of the Democratic two-party vote share in each precinct across the relevant elections, and separately take a geometric mean of the turnout across the elections (due to its skewed distribution), to produce a baseline number of Democratic and Republican votes for each precinct. We then adjust these baseline estimates to balance the total number of Democratic and Republican votes nationwide, reflecting an assumption that in equilibrium the parties will on average each win half the votes.

For example, the pre-adjustment baseline Democratic vote count estimate for precinct j in 2020, denoted by $\hat{D}_j$, can be written as

$$\hat{D}_j = \frac{1}{2}\left(\frac{D_{16j}}{D_{16j} + R_{16j}} + \frac{D_{20j}}{D_{20j} + R_{20j}}\right) \times \sqrt{(D_{16j} + R_{16j})(D_{20j} + R_{20j})}$$

where D_{tj}, R_{tj} are the Democratic and Republican vote counts for the precinct in year t. A geometric mean for turnout values corresponds to the usual mean for log-turnout values.

To perform the nationwide adjustment, we logit-transform each precinct vote share, add a constant shift, then transform back to the original scale. The shift amount is determined numerically so that $\sum_j \hat{D}_j = \sum_j \hat{R}_j$, where j runs over all precincts in the nation.

We model the two-party Democratic vote share in each district y_{it} for an election t as

$$\operatorname{logit}(y_{it}) = \alpha_i + \beta_t + \varepsilon_{it}, \tag{1}$$

$$\beta_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\beta^2),$$
$$\varepsilon_{it} \overset{iid}{\sim} \mathrm{t}_\nu(0, \sigma_\varepsilon^2),$$

where α_i is the district's baseline partisanship, β_t is election-to-election national swing, and ε_{it} is a district-election specific error. This is the Stochastic Uniform Partisan Swing model of Gelman and King (1994), but put onto a logit scale.

Estimating this model poses challenges because of data limitations. In particular, each different simulated plan has its own set of α_i which must be estimated. However, since these plans are hypothetical, we have no data on elections conducted under these plans. So we fix α_i for each district in the enacted plan and the 5,000 simulated plans by simply aggregating our estimate of the precinct-level baseline vote counts for the district. Specifically, for a given plan p, we compute

$$\hat{\alpha}_{ip} = \frac{\sum_{j \in \mathcal{J}_{ip}} \hat{D}_j}{\sum_{j \in \mathcal{J}_{ip}} \hat{D}_j + \sum_{j \in \mathcal{J}_{ip}} \hat{R}_j}$$

where $\mathcal{J}_{ip}$ indicates the set of precincts that are assigned to district i in a redistricting plan p.

Additionally, since we use the model to predict *future* elections, we will not know β_t and ε_{it}. Instead, these are drawn from the normal distributions with the variance of σ_β^2 and t-distribution with ν degrees of freedom and scale σ_ε^2, respectively, under this model. This injects the appropriate amount of uncertainty about future national and district-specific election swings into our election predictions, which are then propagated to uncertainty in our topline estimates that are presented as figures in the main text.

Thus, to create election predictions, it remains to estimate σ_β^2, σ_ε^2, and ν; once these are estimated along with our baseline estimates $\hat{\alpha}_i$, we can simulate hypothetical future election outcomes. To estimate σ_β^2, σ_ε^2, and ν, we fit the model given in (1) to historical House elections. The data from MIT Election Data and Science Lab (2017) contains almost all House elections since 1976. We study only the races contested by exactly one candidate from each party.

In fitting this model, we are constrained by the lack of historical presidential election data disaggregated to the congressional district level. This means that we cannot create estimates of $\hat{\alpha}_i$ in the manner described above that we use for our future predictions. Instead, in the historical election model only, we fit α_i as a random effect, which is specific to each district but constant across elections. To account for redistricting, which changes the districts every decade, we estimate a separate α_i for each district-decade combination (for example, WA–07 from 2012–2020 would receive a single random intercept) as a random effect. None of the α_i estimated as part of our historical House election model are used in the predictions of future elections. The use of random effects here is only to properly allocate the total variability in election returns to three sources: district-specific, year-specific, and district-year-specific effects. Only the estimates of σ_β^2, σ_ε^2, and ν are used to produce the results in the main text.

In principle, it is simple to compute the average number of seats Democrats will win under a particular redistricting plan. We can simulate new vote shares for each district from the fitted election model. Then for each simulation draw, we can compute the number of seats the Democrats win. Averaging this number across simulation draws yields an estimate of the average Democratic seats.

However, this approach introduces Monte Carlo error controlled by the number of draws from the election model. When the number of Democratic seats must be estimated for 50 states and nationwide, not just for the enacted plans but also for 5,000 simulated plans, it can prove computationally costly to have more than a few hundred draws from the election model. The Monte Carlo error will then aggregate as we sum across districts and states, and will generally not be negligible compared to the variance from the redistricting simulation.

To address this issue, we can exactly compute the expectation of the number of Democratic seats with Gauss-Hermite quadrature. Specifically, in a state with n districts, and given estimates $\hat{\boldsymbol{\alpha}}$ of the baseline partisanship of each district, we wish to estimate

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n} \mathbf{1}\{y_{it} > \tfrac{1}{2}\} \,\middle|\, \boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}\right] \\
&\quad = \sum_{i=1}^{n} \mathbb{E}[\mathbb{P}(y_{it} > \tfrac{1}{2} \mid \alpha_i = \hat{\alpha}_i, \beta_t) \mid \alpha_i = \hat{\alpha}_i] \\
&\quad = \sum_{i=1}^{n} \mathbb{E}[\hat{F}_\varepsilon(\alpha_i + \beta_t) \mid \alpha_i = \hat{\alpha}_i],
\end{aligned}$$

where the second step follows from the law of iterated expectations, and in the last step $\hat{F}_\varepsilon$ is the CDF of the error term in the fitted election model (which is a plug-in estimate from the historical House election model). Since the final expectation conditions on α_i, it is only averaging over β_t, which is drawn from a $\mathcal{N}(0, \sigma_\beta^2)$ distribution. Thus we can estimate $\mathbb{E}[\hat{F}_\varepsilon(\alpha_i + \beta_t) \mid \alpha_i = \hat{\alpha}_i]$ to high accuracy using Gauss-Hermite quadrature (Golub and Welsch, 1969), and without resorting to random draws from the election model. Specifically, we use an order-6 approximation, which simulation testing showed was more than enough for extremely high accuracy.