The Curse of the Silent Majority*

Marco Battaglini † Thomas R. Palfrey ‡ Gabriel Sekeres §

November 4, 2026

Abstract

We study a model of elections in which some voters belong to social groups, such as churches, clubs, unions, and social-media networks, through which they can coordinate turnout, while unaffiliated voters decide independently. We identify weak conditions under which unaffiliated voters find it strictly optimal to abstain, leaving the electoral outcome to organized groups. We call this phenomenon the "Curse of the Silent Majority," because even when unaffiliated voters constitute a majority and have stronger preferences, the candidate supported by the organized minority wins with high probability. We show theoretically that this form of group voting can account for empirical regularities that are difficult to explain using models of individual-based turnout decisions, particularly the frequent occurrence of large margins of victory, including in high-turnout elections. We then assess the quantitative relevance of this mechanism by structurally estimating the model using data from Texas liquor referenda. The estimates underscore the importance of organized groups in explaining voter turnout and electoral margins and provide new insights into the social organizations that shape electoral outcomes.

JEL Classification: C72, D72
Keywords: Voter turnout; Electoral Equilibrium; Group-based Voting

*Support from the National Science Foundation (SES-2343948) is gratefully acknowledged. We thank audiences at several seminar and conference presentations for helpful comments and suggestions. Edoardo Bollati provided excellent research assistance.

†Department of Economics, Cornell University, Ithaca, NY 14850. Email: battaglini@cornell.edu

‡Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125. Email: trp@hss.caltech.edu.

§Department of Economics, Cornell University, Ithaca, NY 14850. Email: gs754@cornell.edu

1 Introduction

Elections are the most basic mechanism through which democracies make collective decisions, and thus a core concept in any political economy theory. Yet, our theoretical understanding of voter behavior in elections remains in many respects unsatisfactory. While rational voter models provide a crucial framework for understanding the calculus of voting, they struggle to explain voter participation and certain empirical outcomes. Meanwhile, non-rational behavioral theories offer alternative explanations for voter participation, but they frequently fail to account for well-documented aggregate patterns of strategic behavior and comparative statics.

A key feature of most existing theories of turnout is that they either focus almost exclusively on individual motivations or largely abstract from individual decision-making altogether. The focus on individual motivations is central to so-called *pivotal-voter* models, which cast voters in the tradition of the classic homo economicus. It is also present in most other behavioral theories of turnout, including classic *decision-theoretic models*, *minimax regret* models, and *expressive-voting* models.[1] Even in *rule-utilitarian models*, where voters are guided by ethical considerations and adopt turnout strategies they believe to be in the best interest of their reference group, the emphasis remains on individual motivation. Each voter determines the group-optimal strategy and then implements it independently of the other members. Groups matter only insofar as they provide the reference point for the individual voter's ethical calculus; they do not facilitate communication or provide direct coordination.[2]

At the opposite end of the spectrum, models that account for organized groups emphasize the role of group leaders but simplify the behavior of individual voters (Morton (1987, 1991), Uhlaner (1991), Shachar and Nalebuff (1999)). In these models, leaders exert costly effort to mechanically increase participation within their faction; individual voters, however, are largely passive players. Even more recent work such as Pogorelskiy (2024) and Levine and Mattozzi (2020) that provide new insights on behavior in groups, focus on environments with two large candidate-specific groups in which one leader per candidate sets the standard for all.

What is missing in the theory of turnout in elections is a study of an environment in which the electorate is composed of both a pool of unaffiliated voters (henceforth referred to as "independent" voters) and multiple non-overlapping social groups of affiliated voters.[3] Many voters belong to political parties and unions, or to other tight-knit non-political organizations such as churches, clubs, charities, extended families, social media groups or service-oriented voluntary associations. These

[1] A more detailed discussion of the related literature is presented in Section 1.1 below.

[2] We discuss this feature of rule-utilitarian models in greater detail in Section 1.1 and, especially, in Section 3.2.2.

[3] Our notion of independent voters should not be confused with the usage that refers to voters who are *politically independent* in the sense of not being registered nor usually voting for a specific political party.

organizations facilitate communication and coordination of voting decisions among their members, and thus affect behavior. However, these social groups mostly act independently of each other or are connected with each other only by loose affiliations or sparsely overlapping memberships. For instance, members of a church may be heavily influenced by the content of Sunday sermons, which are shaped by religious doctrine but also by the idiosyncratic taste or political views of the church leader. How does the presence of these social groups affect turnout and the outcome of elections? Can the recognition of these groups help rationalize empirical evidence?

This paper takes a first step in addressing these questions by developing a model of elections in which part of the electorate is comprised of groups that coordinate turnout decisions of their affiliated members, in combination with a residual set of unaffiliated voters who take their individual decisions independently of each other. Voters in groups, as well as independent voters, choose either to vote for their preferred candidate or to abstain, and voting is costly. The goal is to understand the mechanism by which the presence of organized groups affects turnout, electoral outcomes, and overall welfare. We then structurally estimate the model to assess how far this coordination mechanism can reconcile the evidence on turnout and electoral margins.

We present two sets of results. First, we study how voting by groups interacts with voting by individuals. We show that the presence of groups may have oversized effects on the outcome of elections, even if the share of the electorate in groups is small and individuals are more motivated than group members. Specifically, we characterize weak conditions under which we either have equilibria, or we only have equilibria in which independent voters find it optimal to abstain, leading to an election decided exclusively by organized groups affiliated to each candidate. A sufficient condition for the complete crowding out of independents in *all* equilibria is that the intensity of preference for electing their respective candidates is not too dissimilar between independents and organized groups. A sufficient condition for the existence of such an equilibrium is that independents' valuation of the electoral outcome be smaller than the sum of the valuations of the two classes of organized groups; when the two classes of groups have similar valuations, this permits independents' valuation to approach twice either group's valuation. We call this phenomenon, in which independents rationally find it optimal to abstain and leave the election to organized groups, the *Curse of the Silent Majority* (henceforth, CSM), because even when unaffiliated voters constitute a majority and have stronger preferences, the candidate supported by the organized minority wins with high probability.

In the second set of results, we show that the presence of social groups may help reconcile strategic voting models with empirical evidence on elections, which as said is often seen as conflicting with them. As first highlighted by Coate et al. (2008), the issue with these models is

not that they cannot be calibrated to predict the observed expected turnout rates: it is that if they manage to fit expected turnout, then they generate implausible distributions of the margin of victory. For instance, using data on the Texas liquor referenda, Coate et al. (2008) find that margins of victory exceeding 10% or even 20% are common, even in large jurisdictions. Such outcomes are, however, predicted with essentially zero probability by standard strategic voting models. Conditional on election characteristics, the independent-voter benchmarks considered here predict a standard deviation of the signed vote-count margin no greater than $\sqrt{\mu}$, the square root of expected turnout. In elections with high expected turnout, these benchmarks therefore predict a small standard-deviation-to-expected-turnout ratio, $\sigma_{MV}/\mu \leq 1/\sqrt{\mu}$, and hence a distribution of the signed vote-count margin that is highly concentrated around its conditional expectation, relative to expected turnout. This concentration can be substantially weaker in our model with organized groups, where coordinated voting can sustain larger tail probabilities in finite samples.[4]

In Section 4.2, we confront our model with groups with the evidence by structurally estimating it using the same data on Texas liquor referenda as in Coate et al. (2008) and in other recent studies on turnout, which we have updated and extended. The estimation shows that allowing for organized groups substantially improves the model's fit and yields better out-of-sample predictions than the benchmark without groups, although the model still understates large margins of victory in large jurisdictions. Because the model with groups nests the model without groups as a special case, we can formally reject this restriction within our maintained specification, even after accounting for the additional parameters. Although neither the cost-to-value ratios governing voting incentives nor the effective group-size parameter is directly observed, both are identified under the maintained structural model. The estimates therefore provide new insights into the nature of the social organizations that influence electoral outcomes. The model can also be readily compared with alternative non-nested models. By combining strategic participation and within-group coordination, the model provides a unified account of mean turnout rates, their decline with electorate size, and the dispersion of electoral margins, rather than simply introducing over-dispersion into vote counts.

Taken together, these results have implications for how political systems are evaluated and estimated. The tendency has been to focus on "hard" features of political institutions, for instance the electoral rule, and thus pay less attention to other "soft" features, such as the presence, strength,

[4]In this discussion, we abstract from aggregate shocks to voters' preferences, costs, or information, which are absent from the basic versions of all the voting models mentioned above. Observed aggregate shocks do not qualitatively alter the predictions; they simply make the results contingent on their realizations. As first highlighted by Battaglini et al. (2010), however, unobserved aggregate shocks may also help rationalize large margins of victory in elections with high turnout, although under plausible assumptions they are unlikely to account fully for the magnitudes observed in the data. We discuss this additional channel in greater detail below, as it may nevertheless help explain the evidence on margins of victory.

and organization of political factions, or even organized nonpolitical groups. The results presented above, show that even in the simplest system, with two candidates and winner-take-all electoral rule, the electoral outcome critically depends on organized groups.

A limitation of our analysis is that, given the complexity of studying elections with heterogeneous groups, our model is, though more general than the standard pivotal model, necessarily very stylized. First, behavior in groups is modeled by simply assuming that each group adopts a simple correlated strategy for its members. Second, we assume that all groups in each electoral district have the same size m: so we effectively have only groups of size m, and individuals. This allows us to model non trivial interactions between groups and individuals and thus generalize the canonical model in the simplest possible way, but it is restrictive. A more satisfactory approach would allow groups to differ in size, incorporate more nuanced forms of within-group coordination and communication, and permit voters to belong to multiple groups simultaneously. In Section 6.1, we argue that allowing for heterogeneous group sizes does not eliminate the key mechanism underlying our results and can potentially be studied in practical ways by numerical methods. The advantage of our approach is that it allows us to obtain clean theoretical results that then guide our empirical analysis.

1.1 Related literature

The discussion above already noted that models of voter turnout have focused on the individual level without reference to social groups. Classic approaches to formalizing voters' behavior include the decision-theoretic models (see Downs (1957), Tullock (1967), and Riker and Ordeshook (1968)), the pivotal-voter models (see Ledyard (1981, 1984) and Palfrey and Rosenthal (1983, 1985)), the minimax-regret model (Ferejohn and Fiorina (1974)), and the expressive-voting models (Brennan and Lomasky (1993), Hamlin and Jennings (2011)). Although these contributions offer alternative accounts of the individual voter's calculus, differing in strategic sophistication, attitudes toward uncertainty, and objectives, they leave little role for social groups as venues for coordination and information sharing.[5]

A class of individualistic models that is especially relevant for our work is that of rule-utilitarian models (Feddersen and Sandroni (2006) Harsanyi (1977), Coate and Conlin (2004)). These models

[5]In decision-theoretic models, voters are rational and instrumental: they mainly care about the election outcome, but do not fully account for other voters' behavior. In pivotal-voter models, elections are formal games and voters are strategic. Minimax-regret models also feature strategic voters, but only in a weak sense, since turnout decisions are independent of the probability of being pivotal - which is the key endogenous variable in the strategic calculus of voting. In expressive-voting models, voters derive a direct positive utility from the act of voting for a particular candidate. Alternatively there could be psychological costs to the act of non-voting. Dellavigna et al. (2017) offer evidence from a field experiment suggesting that such costs to non-voting can result from a combination of social-image concerns and lying aversion. See Dhillon and Peralta (2002) and Feddersen (2004) for more detailed discussions of turnout models.

are closer to ours because they introduce group thinking into the motivation of each individual voter: a rule-utilitarian voter chooses the strategy that, if adopted by all, would generate maximal utility for the relevant reference group.[6] Still, they can be seen as individualistic because groups do not allow players to coordinate their actions or share relevant information: group utility maximization provides a motivation for a given individual strategy, which is then implemented independently by its members. Our model differs from these models in two ways. First, in our model, a group coordinates its members' actions by having them vote or abstain as a block; it is not merely a reference point for the independent strategies of its members. Second, we assume that there are many groups, interpreted as smaller social units that are often non-political, such as churches, families, unions, and similar associations. Assuming small groups is natural if groups are to facilitate communication and coordination, since asymmetric information and limited trust would hinder both in large, anonymous groups. By contrast, rule-utilitarian models typically involve broad constituencies, each aligned with a candidate.

At the other end of the spectrum are group-based models that focus on leaders rather than voters (Morton (1987, 1991), Uhlaner (1991), Shachar and Nalebuff (1999)). In these models, leaders can mobilize members at personal cost, so elections become games between leaders. Herrera and Martinelli (2006) extend these early group models by endogenizing the number of leaders on each side, each of whom motivates a random number of voters to turn out for their side, and show that the relationship between expected turnout and margin of victory depends on leadership costs and the strength of social interactions. Our model shares with these approaches an emphasis on the importance of groups for coordinating turnout. As in rule-utilitarian models, traditional group-based models typically assume that voters are grouped by candidate, so the strategic interaction occurs between candidate-based "macro" groups or leaders. In our model, groups are smaller entities that strategically interact with one another and with independent voters, both pro and against their favorite candidate.

A few recent contributions, situated between the previous two classes of models, propose theories of social groups in voting that are related to our contribution. Levine and Mattozzi (2020) present a model à la Feddersen and Sandroni (2006), but in which each candidate designs the social norm as an incentive scheme for his or her voters.[7] As in previous rule-utilitarian models, given

[6] In Harsanyi (1977) voters have the same preferences and associate ethical behavior with the strategy that maximizes the welfare of the entire society if adopted by all voters. In Feddersen and Sandroni (2006) and Coate and Conlin (2004) voters are divided in two reference groups, with each associated to a candidate, and they take the strategy of the opposing reference group as given. In Feddersen and Sandroni (2006) ethical voting behavior is that which maximizes the welfare of society; in Coate and Conlin (2004) ethical voting behavior maximizes the welfare of their reference group.

[7] Another related contribution is Levine and Martinelli (2024), which also models elections as a competition between incentive schemes associated to the two candidates. In their paper, leaders motivate followers through peer

the utilitarian rule, voters act independently, but now the social norm is endogenous and part of the candidate strategies. Pogorelskiy (2024) considers a model of voting in which voters can communicate and therefore play correlated equilibria. This is related to our approach in emphasizing coordination among voters, although our model restricts coordination within each group to simple block strategies. Pogorelskiy (2024) characterizes the set of correlated equilibria of a voting game with two candidates under alternative informational assumptions, focusing on the equilibrium with the highest turnout. As we will show, these theories generate empirical predictions that differ substantially from ours.

From a more empirical perspective, structural estimates of voting models are presented by Shachar and Nalebuff (1999), Coate and Conlin (2004), Coate et al. (2008), Kawai and Watanabe (2013), Kawai et al. (2021) and Berry et al. (2026), among others.[8] Shachar and Nalebuff (1999) show that turnout in U.S. presidential elections is correlated with expected closeness, a finding they view as difficult to reconcile with expressive-voting models, where closeness plays no role, and with pivotal-voter models, where pivotal probabilities are uniformly small. They structurally estimate the leadership model described above and show that it can rationalize this evidence. They do not, however, compare its performance with alternative models such as the pivotal-voter model, nor do they study whether it can explain the overall distributions of turnout and margins of victory. Coate and Conlin (2004) structurally estimate a version of the rule-utilitarian model and show that it can explain turnout levels. They show that the model performs better than an expressive-voting model, but they do not compare it with a pivotal-voter model, nor do they examine whether it can account for other key statistics, such as the margin of victory. Coate et al. (2008) is the closest paper to ours: it explicitly studies the pivotal-voter model, which is a special case of our framework, and compares it with the expressive-voting model. They are the first to make the key observation that pivotal-voter models fit turnout surprisingly well, but fail to explain the margin of victory. Kawai and Watanabe (2013) structurally estimates a model inspired by Myerson and Weber (1993) using Japanese general-election data. They find that 63 to 85 percent of voters vote strategically, but, unlike our model and the other models discussed above, their focus is on candidate choice in multi-candidate elections rather than on turnout. Kawai et al. (2021) and Berry et al. (2026) differ from both the earlier literature and our work in that they place less emphasis on specifying and identifying a theoretical model. Instead, they study how voter turnout affects the aggregation

pressure and costly monitoring, but also by keeping the election close, thereby maintaining a sufficiently high pivot probability.

[8]For reasons of space, we focus here on structural estimates of voting models using field data. There is also a rich experimental literature; see Hansen et al. (1987), Levine and Palfrey (2007), Battaglini et al. (2008, 2010), and Palfrey and Pogorelskiy (2019), among others. Insightful reduced form empirical analyses are presented, for instance, by Fujiwara (2011) and Bursztyn et al. (2024), among others.

of preferences in U.S. presidential and congressional elections, relying on reduced-form models in which pivot probabilities are specified as functions of demographic characteristics and exploiting rich data on voter behavior. We discuss these approaches in greater detail in Section 5.2.3, where we compare their performance with that of our model.

2 Model

2.1 Setup

Consider an election with two alternatives L and R. There are two types of voters: independent voters and organized voters who are affiliated with a group. Independent voters are the 'standard' type of voters typically assumed in traditional voter turnout models (Ledyard (1981, 1984) and Palfrey and Rosenthal (1983, 1985)), each of whom rationally and independently selects whether to vote or abstain, and who to vote for on the basis of their expectations on how other voters vote and their preferences. Following the large games approach of Myerson (1998, 2000), we assume there are $\widetilde{N}_I$ independent voters, where $\widetilde{N}_I$ is a Poisson distributed random variable with mean N_I. The independent voters are assumed to receive utility $v_I > 0$ if their favorite candidate is elected, and 0 otherwise.

In addition to the independents, there are organized groups of voters. These groups may correspond to chapters of organized societies such as churches, sports clubs, unions, or other organizations such as the National Rifle Association or even local party organizations. Group membership is treated as exogenous in the baseline model. Members can, however, leave their group and choose their voting action independently while retaining their preferences. In Section 6.2 we verify that remaining in the group is individually rational in the equilibria of interest and discuss the incentives for compliance with the group's recommendation.

We assume that there are $\widetilde{J}_j$ associated groups in correspondence to each candidate $j = L, R$, each group with an integer number $m \geq 2$ of members, where $\widetilde{J}_j$ is a Poisson distributed random variable with mean $J_{j,N}$. We assume that $\widetilde{N}_I$, $\widetilde{J}_L$, and $\widetilde{J}_R$ are mutually independent. Voters and groups know the parameters of the model but not the realizations of the numbers of other voters and groups when they choose their actions. The expected number of voters in groups of type j is therefore $N_j = mJ_{j,N}$. Similarly to independent voters, each citizen in a group $j = L, R$ receives a utility equal to v_j from the election of the favorite candidate $j = L, R$, and a utility of zero from the election of the other candidate. We assume that groups act independently of each other. The idea is that there are natural limits to coordination across groups. As much as a national club or society may desire to coordinate the behavior of its chapters, local groups remain independent and coordinate at the local level only. If the election is tied, each alternative wins with probability $1/2$.

Voting has a cost $c > 0$ for each voter, both for individual voters and for voters in groups. As standard in the literature on voting, we assume:

$$c/v_i \in (0, 1/2)$$

for $i = L, R, I$. This implies that expected participation in an election cannot be zero in equilibrium: if a voter expects no turnout (and so a tie), then s/he prefers to vote.

Independent voters and the voters in groups of type j constitute, respectively, a share α_I and α_j of the total size of the expected electorate such that $\alpha_I + \alpha_L + \alpha_R = 1$, which is denoted N (so $N_l = \alpha_l N$ for $l = I, L, R$). We assume that $\alpha_I, \alpha_L, \alpha_R > 0$. In the large-election results of Sections 2.2 and 3.1, the population shares and group size are held fixed as N increases. As a leading case, we assume that independents are the largest group, so $\alpha_I > \alpha_R, \alpha_L$ and, for simplicity, that they all prefer L to R.[9] It follows that the voters who support L constitute a share $1 - \alpha_R > 1/2$ of the electorate: so L is the ex ante most popular candidate.

To reduce the number of cases to study and focus the analysis on the most interesting cases, we make two additional assumptions. First, our goal is to show how the presence of organized groups, even if they comprise a small fraction of society and they are less motivated than independents, can induce abstention by *all* the independents and leave elections lopsided. Therefore, we initially assume independents are more motivated than groups, so $v_I > \max\{v_L, v_R\}$. Doing so stacks the odds against abstention by independents. The second assumption concerns the group's relative preferences. Since L is the most popular candidate, it is natural to contrast it with a candidate who has the most intense support from its groups, so we assume $v_R > v_L$. This, as we will see, guarantees that there are equilibria in which the minority candidate R prevails if only groups voted in the elections. These assumptions allow us to focus on the cases of greatest interest and make the analysis more transparent. In particular, they provide a crisp benchmark against which to assess the implications of independents' equilibrium abstention.

As is standard in the literature on voting, voters and groups avoid weakly dominated strategies, so they either abstain or vote for their favorite candidate. An independent voter i's strategy is $\sigma_i \in [0, 1]$, where σ_i is interpreted as the probability i votes for L, and $1 - \sigma_i$ is the probability of abstaining. We focus on symmetric strategies, where each independent voter votes with probability σ.

The coordinated voting strategy of the voters who belong to a group is potentially a more complex object. A member of a group does not independently decide to vote; the group coordinates

[9]The assumptions that $\alpha_I > \max\{\alpha_R, \alpha_L\}$ and that all independent voters prefer L simplify the benchmark analysis and the characterization of equilibria in which independents participate by reducing the number of cases to consider. The existence results below with groups in which all independents abstain, which constitute the core of our paper, do not depend on the composition of independents' political preferences.

its members' turnout decisions. In the literature on group based ethical voting started by Feddersen and Sandroni (2006), groups exist as references to assess ethical decisions. Voters choose strategies following an ethical criterion: voter i in a group j chooses the probability σ_i^j that if adopted by all members of j, maximizes the payoffs of group j's members. In this approach, the strategy space of a group member is exactly the same as that of an individual, a probability in $[0,1]$; what differs is the criterion with which σ_i^j is chosen. We instead assume that groups are organizations that coordinate the turnout decisions of their members.[10] The idea is that organizations such as civic associations or churches provide a setting in which members can communicate and coordinate their turnout decisions. [11] A group chooses its action to maximize the sum of its members' expected utilities. Since all members of a type-j group have identical preferences, this is equivalent to maximizing the expected utility of a representative member. We assume that each group adopts a simple strategy under which all its members either abstain or vote as a block. We restrict attention to strategies that are symmetric across groups of the same type. These strategies treat all group members symmetrically ex post and require only a common public recommendation; incentives for compliance are examined in Section 6.2. A *simple strategy* for a type-j group can therefore be described by a number $\beta_j \in [0,1]$, where β_j is the probability that all m members vote for their favorite candidate, and $1-\beta_j$ is the probability that all members abstain. The restriction to block strategies simplifies the analysis, as we do not seek to model the details of groups' internal deliberations. It can be interpreted as an organizational constraint in groups whose coordination is limited to uniform participation obligations rather than individualized assignments or recommendations for independent randomization. With this definition, a quasi-symmetric voting equilibrium with groups is defined as follows:[12]

Definition 1. *A profile (σ, β) is a voting equilibrium with groups if:*

- *An independent voter's probability of voting σ is optimal given the strategy profile (σ, β);*
- *A group of type j's probability of voting, β_j, for $j = L, R$, is optimal for the group given the strategy profile (σ, β).*

In what follows, we refer to a quasi-symmetric voting equilibrium simply as an equilibrium.

[10] The groups act independently of each other, so the coordination occurs within each group, not across groups.

[11] As in Dixit (2000), we assume in the baseline model that group members comply with the group's recommendations. Section 6.2 provides a micro-foundation for this assumption by studying members' incentives to comply.

[12] The existence of a quasi-symmetric equilibrium of the game is guaranteed by standard arguments.

2.2 A Benchmark: Voting equilibrium without group coordination ($m = 1$)

We first consider a benchmark environment with the same expected demographic composition as above, but in which all voters, including those associated with a group, vote independently. In other words, in this benchmark case we characterize the voting equilibria as in Definition 1, for the special case of $m = 1$. The expected number of independents (who all prefer L and have a value v_I) is $\alpha_I N$, the expected number of L-group members, each of whom has a value of v_L, is equal to $\alpha_L N$, and the expected number of R-group members, each of whom has a value of v_R, is equal to $\alpha_R N$. Recall that $v_I > v_R > v_L$ and the voting cost c is the same for all voters.

Voter turnout games can have multiple quasi-symmetric equilibria. We first establish that, even if there are many voting equilibria without groups, all such equilibria share two key properties that allow us to pin down the outcome of the election. We say that a property holds "in a large election" if, given all the other parameters, there is a $\overline{N}$ such that the property holds for all $N > \overline{N}$. We have:

Lemma 1. *In a large election without organized group coordination, $\sigma \in (0, 1)$ and $\beta_L = 0$ in any equilibrium. That is, independents vote for L with a strictly positive probability less than 1, and L-type group members always abstain.*

Proof. See Appendix. □

The logic behind Lemma 1 is straightforward. For N large, there cannot be an equilibrium where independents vote for their favorite candidate L with probability one, since in this case L would win with probability approaching one and their probability of being pivotal would be too low. At the same time, independents cannot abstain. Since $v_I > v_L$ and independents and L-types are pivotal in the same events in the absence of organized groups, this would imply that L-types abstain as well. But if only R-types vote, an independent is pivotal when zero or one R-type votes, whereas an R-type is pivotal only when no other R-type votes. Together with $v_I > v_R$, this rules out such an equilibrium.

We now use Lemma 1 to characterize the outcome of the election when c is small. In general, even if N is large, the probability of success for either candidate is interior in equilibrium, so the most popular candidate, L, loses with positive probability. This occurs because with a finite c, turnout must be bounded even if the electorate is arbitrarily large. If turnout were too large, the probability of being pivotal for each voter would be too small to make participation incentive compatible. When c is small and N is large, however, the probability of success for L should be expected to be large. We can use Lemma 1 to show that this is indeed what must happen in

equilibrium. For any sequence of equilibria indexed by N, let $P_N^L(c)$ denote the probability that L wins along the sequence and define $P^L(c) = \liminf_{N\to\infty} P_N^L(c)$. Proposition 1 applies to every such sequence. We have:

Proposition 1. *For any $\delta > 0$, there is a $c(\delta)$ such that for $c < c(\delta)$, we have $P^L(c) \geq 1 - \delta$.*

Proof. See Appendix. □

Proposition 1 implies that, for sufficiently small c, the probability that L wins is arbitrarily close to one in large elections along every equilibrium sequence. This is the benchmark against which we will compare what happens in the presence of organized groups.

The result relies on the assumption that $v_I > v_R$. When $v_I \leq v_R$, the comparison is less clear-cut because R-types are at least as motivated as independents, and the electoral outcome also depends on their population share α_R. The assumption $v_I > v_R$ therefore provides a particularly clear benchmark for the analysis with groups.

3 Equilibrium with organized groups

In Section 3.1, we prove that under general conditions, in the presence of organized groups, independents abstain, leaving the outcome to the groups. In Section 3.2, we prove that the prevalence of organized groups can help explain empirical evidence on voting behavior, resolving fundamental issues well known in the literature.

3.1 The independent voters curse

The previous section established that in large elections when c is small (not necessarily infinitesimal), independent voters will propel L to victory with high probability in all equilibria. Indeed, as $c \to 0$ the limit probability of success always converges to 1 in equilibrium. In this section we study electoral competition in the presence of organized groups and show that results are very different even though the groups may comprise an arbitrarily small fraction of the population and are strictly less motivated than the independents (i.e., v_L and v_R are strictly smaller than v_I).

Proposition 2 below derives a simple sufficient condition for the existence of an equilibrium in which all independents abstain, the election is completely determined by the groups, and party R wins with probability strictly greater than $1/2$, approaching probability 1 as $c \to 0$.

The strategic logic for why even an arbitrarily small share of organized voters can upend an election is rather subtle, and can be explained as follows. When independents abstain, in a large election the groups vote for their favorite candidate with interior probabilities β_R and β_L. These

probabilities solve the system:

$$c = \frac{v_R e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2\frac{N}{m}}}{4\left(\pi\frac{N}{m}\sqrt{\alpha_R\beta_R\cdot\alpha_L\beta_L}\right)^{\frac{1}{2}}}\left(1+\frac{(\alpha_L\beta_L)^{\frac{1}{2}}}{(\alpha_R\beta_R)^{\frac{1}{2}}}\right) \quad (1)$$

$$c = \frac{v_L e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2\frac{N}{m}}}{4\left(\pi\frac{N}{m}\sqrt{\alpha_R\beta_R\cdot\alpha_L\beta_L}\right)^{\frac{1}{2}}}\left(1+\frac{(\alpha_R\beta_R)^{\frac{1}{2}}}{(\alpha_L\beta_L)^{\frac{1}{2}}}\right) \quad (2)$$

The left hand side of each of these equations is the cost of voting; the right-hand side is the value of electing the preferred candidate times the pivot probability, that is the probability that the vote by a group changes the election toward its favored candidate.[13] An R group (resp., L group) is pivotal in two events: first, when the two parties are tied (the first term in the right-hand side of (1)), that is when the number of R groups voting equals the number of L groups voting, in which case an R group will break the tie by voting; second, when the number of R groups voting is exactly one less than the number of L groups voting (the second term in the right-hand side of (1)), in which case an R group will create a tie by voting. Equation (2) applies to the L group similarly. In equilibrium, both groups vote with interior probability only when the right-hand side of equation (1) equals the right-hand side of equation (2).

Therefore, a key step is to establish that, in large elections, when the groups satisfy equations (1) and (2), which assumes abstention by independent voters, it is indeed optimal for the independents to abstain even if v_I is strictly larger than v_R and v_L. This requires the following condition to be satisfied:

$$c \geq \frac{v_I e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2\frac{N}{m}}}{4\left(\pi\frac{N}{m}\sqrt{\alpha_R\beta_R\cdot\alpha_L\beta_L}\right)^{\frac{1}{2}}} \quad (3)$$

where the right-hand side is v_I times the equilibrium pivot probability for an independent. The pivot probability for an independent is different than that of the groups because an independent is pivotal only when the groups' votes are tied. If L is one group below R, an independent cannot do anything to change the outcome of the election, since R has m votes more than L.

Conditions (1) and (2) have two implications. The first implication is obtained by equating the RHS of equations (1) and (2). The exponential term and the denominator cancel out, implying the equilibrium condition:

$$\frac{\alpha_R\beta_R}{\alpha_L\beta_L} = \left(\frac{v_R}{v_L}\right)^2. \quad (4)$$

[13] We write these pivot probabilities using the approximations developed in Myerson (2000), which converge to the true values in elections with large participation. Here large participation refers to the expected numbers of participating groups, which need not increase with N when $c > 0$ is held fixed. The actual proof of Proposition 2 does not use approximations and proves the result in full generality. Re-estimating the full-sample grouped specification using the exact Poisson pivotal probabilities in place of these approximations, under the same fixed mean-zero Skellam noise rule, yields virtually identical estimates: every fitted integer group size is unchanged, and the mean absolute change in predicted turnout is 0.033 percentage points.

This condition lends a simple and intuitive interpretation. If, for instance, $v_R > v_L$, then $\alpha_R\beta_R > \alpha_L\beta_L$: expected turnout for R is greater than expected turnout for L, implying that the pivot probability for an R group is *less* than the pivot probability for an L. This is necessary in order to satisfy (1) and (2). Since the groups have different values, there must be a smaller pivot probability for the more motivated group, and a higher pivot probability for the less motivated group. Condition (4) defines the exact equilibrium relationship between β_R and β_L in order for (1) and (2) to be simultaneously satisfied.

The second implication is obtained by dividing equations (1) and (2) by v_R and v_L, respectively and then subtracting (1) from (2):

$$\begin{aligned} \frac{c}{v_L} - \frac{c}{v_R} &= \frac{e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2\frac{N}{m}}}{4\left(\pi\frac{N}{m}\sqrt{\alpha_R\beta_R\cdot\alpha_L\beta_L}\right)^{\frac{1}{2}}}\left(\frac{(\alpha_R\beta_R)^{\frac{1}{2}}}{(\alpha_L\beta_L)^{\frac{1}{2}}} - \frac{(\alpha_L\beta_L)^{\frac{1}{2}}}{(\alpha_R\beta_R)^{\frac{1}{2}}}\right) \\ &= \frac{e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2\frac{N}{m}}}{4\left(\pi\frac{N}{m}\sqrt{\alpha_R\beta_R\cdot\alpha_L\beta_L}\right)^{\frac{1}{2}}}\left(\frac{v_R^2 - v_L^2}{v_L v_R}\right) \Leftrightarrow \frac{e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2\frac{N}{m}}}{4\left(\pi\frac{N}{m}\sqrt{\alpha_R\beta_R\cdot\alpha_L\beta_L}\right)^{\frac{1}{2}}} = \frac{c}{v_R+v_L} \end{aligned} \tag{5}$$

where in the second equality we used (4). Combining this condition with (3), we obtain the condition:

$$\frac{c}{v_I} \geq \frac{c}{v_R+v_L} \Leftrightarrow v_I \leq v_R + v_L \tag{6}$$

It follows that when (6) holds, then it is an equilibrium for the groups to vote with probabilities β_R and β_L and for independents to abstain completely. There are two important qualitative properties of the equilibrium. First, (4) implies that the expected turnout for R is strictly greater than expected turnout for L, so R wins with probability strictly greater than 1/2. Second, the result holds for any value of $\alpha_j > 0$ for $j = L, R$. Thus, even if the size of population organized as groups is arbitrarily small ($\alpha_j \approx 0$ for $j = L, R$), and even though independents have the highest value for electing their candidate, they find it optimal to abstain in the equilibrium with organized groups. Indeed, in the empirical analysis of Section 5.2.2, we show that the evidence from Texas liquor referenda is consistent with organized groups comprising as little as 10% of the population on each side.

The analysis above was simplified in the sense that it substituted the Myerson (2000) approximation formulas for the pivot probabilities. The following proposition formally states and proves the result using the exact formulas for pivot probabilities. This leads to two minor technical differences in the statement of the result. First, we impose $c < \bar{c}$, a sufficient condition ensuring that, for sufficiently large N, both types of groups mix with interior probability in the auxiliary game in which independents abstain. Second, rather than $v_I \leq v_R + v_L$ as in the limit case, the upper

bound on v_I is $v_R[1+\Delta^*]$ for some $\Delta^* > 0$, that is independent of N. Define

$$\bar{c} = (v_R/2) \cdot e^{1-v_R/v_L}$$

Note that if $v_R = v_L$ then $\bar{c} = v_R/2$. We have:

Proposition 2. *Assume $c < \bar{c}$. There exists $\Delta^* > 0$ such that if $v_I \leq v_R[1+\Delta^*]$ then in a large election with organized groups there is an equilibrium in which $\sigma = 0$, $\beta_R, \beta_L \in (0,1)$, and $\alpha_R\beta_R > \alpha_L\beta_L$. At this equilibrium, the probability that L wins is strictly less than $1/2$.*

Proof. See Appendix. □

The main driver of the result is that since groups vote as coordinated blocks, they are pivotal in more events than individuals, who are pivotal only in the event of a tie. In a large election, the difference in pivot probabilities may become very small, but it is always proportional to $\frac{c}{v_R+v_L}$, so it does not become negligible relative to c/v_I even if N is arbitrarily large.

The condition for Proposition 2 to hold is that v_I is not too large, $v_I \leq v_R[1+\Delta^*]$. This is natural since if the independents have an extremely large value for electing L, then they would always find it optimal to vote. What is significant is that the result holds even if $v_I > v_R$.

The size of Δ^* naturally depends on equilibrium participation rates of the organized groups, which depend on the ratio c/v_j for $j = L, R$. We obtain a stronger result if c is small:

Proposition 3. *If $v_I < v_R + v_L$ then, for every $\delta > 0$ there is a $c(\delta)$ such that, for every $0 < c < c(\delta)$, every sufficiently large election with organized groups admits an equilibrium in which independents abstain and L wins with probability less than δ.*

Proof. See Appendix. □

Proposition 3 has two lessons. First, the relevant pivotality advantage converges to v_L/v_R as $c \downarrow 0$. When c is small, participation is large in equilibrium and we are essentially in the limit case described above, in which the Myerson (2000) formulas for the pivot probabilities become asymptotically exact. This is important because it shows that, in the plausible case in which members of the R and L groups have similar preference intensities ($v_L \simeq v_R$), then Δ^* is approximately equal to 1. The equilibrium in which independents abstain and become irrelevant, therefore, exists even if independents have a valuation arbitrarily close to twice that of either organized group. The second lesson from Proposition 3, is that in the presence of organized groups, when c is small and N large, there is an equilibrium in which L loses the election with probability close to 1; without groups, Proposition 1 establishes that, for sufficiently small c, L wins with probability arbitrarily close to one in all equilibria of sufficiently large elections.

A final question concerns whether abstention of the unaffiliated voters is necessary. Propositions 2 and 3 obtain a simple sufficient condition under which there *exists* an equilibrium in which the independents do not vote, and the decision is left to the organized groups, leading to the success of R with probability strictly greater than $1/2$; i.e., a likely victory for R is possible in equilibrium, not that it is inevitable. Under what conditions this conclusion is true for all equilibria?

For Proposition 4 only, we relax the baseline assumption $v_I > v_R$, while maintaining $v_I > 0$ and $v_R > v_L$. Define $\vartheta = v_I - v_L$. The following result characterizes a simple and natural condition under which abstention by independents holds for all equilibria:

Proposition 4. *There exist $\vartheta^* > 0$ and $\bar{N} < \infty$ such that, whenever $\vartheta < \vartheta^*$, every equilibrium satisfies $\sigma_N = 0$ for every $N \geq \bar{N}$.*

Proof. See Appendix. □

Note that when ϑ is sufficiently small, $v_I < v_R + v_L$ is automatically satisfied. We maintain $v_R > v_L$ throughout, but the assumption $v_I > v_R$ is not essential for Proposition 4: the proof requires only that v_I not exceed v_L by more than a sufficiently small amount.

3.2 Why Organized Groups Help Reconcile Data with Theory

In this section, we explain why the Curse of the Silent Majority (hence, CSM), that is, the phenomenon according to which the election is dominated by the groups, may help reconcile strategic behavior based on instrumentally motivated voters with empirical evidence that is generally considered at odds with such behavior.

As a preliminary step, in Section 3.2.1 we discuss why and to what extent the model is identified by the data. In Section 3.2.2, we explain the theoretical reasons why the model with only individual citizens and no social groups, so with $m = 1$, cannot fit the existing data. As we will make clear, the problem with the case $m = 1$ is not simply a feature of pivotal-voter models. It is instead a feature of models in which voters act independently of each other, conditional on election characteristics and in the absence of unobserved aggregate shocks or other common sources of correlation, regardless of the specific source of motivation. Thus, our argument for incorporating groups in the theory of voting applies equally well to benchmark versions of behavioral models such as classic decision-theoretic models, expressive-voting models, ethical-voter models, and social-image models.

Finally, in Section 3.2.3 we present the theoretical reason why the presence of groups helps resolve this problem and may reconcile theory with data. The more practical empirical question of the extent to which an estimated model with groups fits the evidence is left to Section 4, where we structurally estimate the model using data on Texas liquor referenda.

3.2.1 Empirical identification of the model

For this identification exercise, when $m = 1$, we focus attention to equilibria in which the most motivated L-supporting class and the R-types participate with interior probabilities. When $m > 1$, we select an equilibrium in which independents abstain and both organized types participate with interior probabilities.

Using the large-participation approximations introduced above, the equilibrium conditions can be approximated by the following two equations:[14]

$$\frac{\exp\left\{-\left(1-\frac{v_L}{v_R}\right)^2 \frac{N_R}{m}\beta_R\right\}}{4\left(\pi\frac{v_L}{v_R}\frac{N_R}{m}\beta_R\right)^{1/2}} = \frac{c}{v_R+v_L} \tag{7}$$

$$\text{and } \beta_L = \frac{N_R\beta_R}{N_L}\left(\frac{v_L}{v_R}\right)^2 \tag{8}$$

where N_i and v_i are the size and value of voters who prefer candidate i. To interpret conditions (7)-(8), note that by Propositions 1 and 2 for all m, the pool of voters for candidate R is constituted by the partisans with value $v_R = v_R$, who have a size $N_R = \alpha_R N$. The pool of voters for candidate L, however, depends on m. When $m = 1$ and so there are no organized groups, under the baseline assumptions, Lemma 1 implies that only independents participate on the L side, so $v_L = v_I$ and $N_L = \alpha_I N$. If instead $v_L > v_I$, the maintained interior equilibrium has only L-types participating on that side, so $v_L = v_L$ and $N_L = \alpha_L N$.[15] When there are groups of size $m > 1$, participation also depends on the values. Throughout the analysis, we assume that the (weak) condition in Proposition 2 and 4 are satisfied and focus on the equilibrium in which only groups vote. In this case, in (7)-(8), we have $v_L = v_L$, and $N_L = \alpha_L N$.

For all of these cases, conditions (7) and (8) give the equilibrium strategies of the groups or individuals β_R, β_L as a solution of a system of two equations.[16] The parameters that determine the equilibrium are c, v_j, α, N, m. This parameter set, however, can be drastically simplified. To this goal, first note that the preference of voters of type i depends only on the ratio $\gamma_i = c/v_i$. The reason is that voter i's net expected benefit from voting is $v_i Piv_i - c$, where Piv_i is the probability that voter i is pivotal. Dividing by v_i, this is equivalent to $Piv_i - \gamma_i$, where $\gamma_i = c/v_i$. It follows that c and v_i affect voting incentives only through the ratio $\gamma_i = c/v_i$, and hence can be identified only through this ratio.

Second, note that, for any $m \geq 1$, α_R and β_R always appear together in (7) and (8) in the form

[14] These conditions are derived from (1), (2) and (4).

[15] For simplicity, we ignore here the non generic case of equal valuations, $v_I = v_L$.

[16] Note that when N is small, we might have corner solutions in which β_i and/or β_{-i} are at a corner, either zero or 1. We ignore these cases here for simplicity.

of $\alpha_R\beta_R$ since $N_R = \alpha_R N$; and when $m > 1$ or $m = 1$ and $v_L > v_I$, α_L and β_L also always appear together in (7) and (8) since $N_L = \alpha_L N$. It follows that when $m > 1$, or when $m = 1$ and $v_L > v_I$, then only the combined variables $\theta_R = \alpha_R\beta_R$ and $\theta_L = \alpha_L\beta_L$ are identified. Instead, when $m = 1$ and $v_L \leq v_I$, only the variables $\theta_R = \alpha_R\beta_R$ and $\theta_L = \alpha_I\beta_L$ is identified.[17] The equilibrium can therefore be characterized and estimated using the limiting indifference conditions:

$$\frac{\exp\left\{-\left(1-\frac{\gamma_R}{\gamma_L}\right)^2 \frac{N}{m}\theta_R\right\}}{4\left(\pi\frac{\gamma_R}{\gamma_L}\frac{N}{m}\theta_R\right)^{1/2}} = \frac{\gamma_L}{1+\gamma_L/\gamma_R} \tag{9}$$

$$\text{and } \theta_L = \theta_R\left(\frac{\gamma_R}{\gamma_L}\right)^2 \tag{10}$$

with the interpretation that $\theta_R = \alpha_R\beta_R$ and $\theta_L = \alpha_L\beta_L$ if $m > 1$ or $m = 1$ and $v_L > v_I$; and $\theta_L = \alpha_I\beta_L$ if $m = 1$ and $v_L \leq v_I$.

To discuss identification under ideal conditions, suppose that, for every pair (N_k, X_k) in the support of the data, we observe an arbitrarily large number of elections with those same characteristics. Here we can take X_k as a list of observable variables such as demographic characteristics, observable median income, etc. This of course will not be the case, since for some combinations of features we will have only few observations, so the estimation will require a more elaborate structural model that will be discussed in Section 4. The discussion of this limit case, however, will help understand the identification of the model in the ideal scenario.

Consider here the case with $m > 1$ (the case with $m = 1$ is analogous). From the data, and each electoral district with observable characteristics N and X, we observe $\widetilde{\mu}_j$ and $\widetilde{\sigma}_j^2$, where $\widetilde{\mu}_j$ and $\widetilde{\sigma}_j^2$ are the empirical mean and variance of the turnout of group j in cities of the same type.[18] We now show that these observations allow us to uniquely identify γ_L, γ_R and m, the unobserved parameters in these districts.[19] To this goal, note that the number of voting groups for candidate j is Poisson, so the mean and variance of turnout are:

$$\mu_j = m\left(\frac{\alpha_j N}{m}\right)\beta_j = N\theta_j \tag{11}$$

$$\sigma_j^2 = m^2\left(\frac{\alpha_j N}{m}\right)\beta_j = mN\theta_j = m\mu_j \tag{12}$$

[17] The reason is that the equilibrium condition sets the probability of voting at a level such that turnout makes groups just indifferent between voting and abstaining. What matters for a group's indifference condition is the level of turnout of the other groups, not whether this turnout is driven by the size of the populations in those groups or by the probability with which they vote. For $m > 1$, Proposition 3 establishes that, if $v_I < v_R + v_L$, then for every sufficiently small $c > 0$ and every N sufficiently large given c, there exists an equilibrium in which independents abstain. In this equilibrium, matching the mean vote count $\widetilde{\mu}_R$ requires $\alpha_R N\beta_R = \widetilde{\mu}_R$, so $\beta_R \leq 1$ implies the lower bound $\alpha_R \geq \widetilde{\mu}_R/N$.

[18] More precisely, electoral district k with observable characteristics N_k and X_k, we observe $\widetilde{\mu}_{j,k}$ and $\widetilde{\sigma}_{j,k}^2$. Since the index k is maintained constant in the discussion, it is omitted in the text for simplicity.

[19] Again, to be precise, for each district k, we obtain $\gamma_{L,k}$, $\gamma_{R,k}$ and m_k.

For given voting strategies, the presence of groups, therefore, does not affect expected turnout, but it affects its variance, which increases linearly in m. Taking the ratio of θ_R to θ_L, equation (10) implies:

$$\frac{\gamma_R}{\gamma_L} = \sqrt{\frac{\widetilde{\mu}_L}{\widetilde{\mu}_R}} \tag{13}$$

We also have that $\widetilde{\sigma}_R^2 + \widetilde{\sigma}_L^2 = m\left(\widetilde{\mu}_R + \widetilde{\mu}_L\right)$. Therefore, we obtain:

$$m = \frac{\widetilde{\sigma}_R^2 + \widetilde{\sigma}_L^2}{\widetilde{\mu}_R + \widetilde{\mu}_L} \tag{14}$$

These simple observations suggest that just observing the mean and variance turnout of the two candidates, we can uniquely identify γ_R/γ_L and m. To verify that we can independently identify γ_L and γ_R, note that we can rewrite (9) using (10), (13) and (14). This gives us:

$$\gamma_R = \frac{\exp\left\{-\left(1-\sqrt{\frac{\widetilde{\mu}_L}{\widetilde{\mu}_R}}\right)^2 \frac{\widetilde{\mu}_R(\widetilde{\mu}_R+\widetilde{\mu}_L)}{\widetilde{\sigma}_R^2+\widetilde{\sigma}_L^2}\right\}}{4\left[\pi\sqrt{\frac{\widetilde{\mu}_L}{\widetilde{\mu}_R}}\frac{\widetilde{\mu}_R(\widetilde{\mu}_R+\widetilde{\mu}_L)}{\widetilde{\sigma}_R^2+\widetilde{\sigma}_L^2}\right]^{1/2}}\left(1+\sqrt{\frac{\widetilde{\mu}_L}{\widetilde{\mu}_R}}\right) \equiv \widetilde{\gamma}_R. \tag{15}$$

which uniquely characterizes γ_R (and thus $\gamma_L = \sqrt{\widetilde{\mu}_R/\widetilde{\mu}_L}\cdot\widetilde{\gamma}_R$).

Identification rests on the maintained model since, as noted above, election-level shocks not captured by the covariates could also generate excess dispersion and large tail probabilities. Our empirical objective is to assess the quantitative explanatory power of this mechanism, rather than to establish it as the unique source of the observed patterns.

3.2.2 The problem with individual-based models of voter turnout

We can now discuss why the standard model with only independent voters has trouble rationalizing the empirical evidence on voting.

The standard critique of individual-based instrumental models of voting (in which each voter acts independently and cares only about affecting the outcome of an election) is that they rely on the assumption that the expected utility of voting is proportional to the pivotal probability, i.e. the probability a single voter determines the outcome of an election. In a large election, the pivotal probability is typically very small, so participation is expected to be low. This is often seen as in conflict with the evidence that turnout is significant even in large elections.

As mentioned in the introduction, in an important contribution Coate et al. (2008) (henceforth, CCM) sharpened the critique by arguing that the main shortcoming of the standard pivotal voter model is not that it cannot account for empirical *average* levels of turnout (T) and margin of victory (MV), but that it fails to explain other key qualitative moments of the distribution of turnout. Specifically, using data from the Texas liquor referenda, Coate et al. (2008) show that a structurally

Table 1: CCM-Sample Turnout

Panel	Elections	Data	CCM	$m = 1$	Grouped
$N < 247$	45	0.630	0.675	0.647	0.615
$247 \leq N < 434$	48	0.553	0.524	0.532	0.554
$N \geq 434$	48	0.432	0.433	0.445	0.439

Notes: All 141 available CCM elections with $0 < N \leq 900$ are retained, including seven with $N < 100$. The fixed bins are $N < 247$, $247 \leq N < 434$, and $N \geq 434$. The CCM column re-estimates the finite pivotal-voter model with the Coate et al. (2008) covariates. Turnout is $(V_R + V_L)/N$.

estimated model predicts turnout reasonably well. The cost of voting is an unobserved variable to be estimated, and plausible distributions of these costs are compatible with the evidence.[20] The estimated model, however, fails to predict the Margin of victory by a significant margin.

Tables 1 and 2 illustrate this point. Table 1 shows the average turnout in the Texas liquor referenda for small ($N < 247$), medium ($247 \leq N < 434$), and large ($434 \leq N \leq 900$) electoral districts, containing 45, 48, and 48 elections, respectively. The table compares these observations with predictions from our re-estimation of the CCM pivotal-voter model and our model, discussed in Section 4.[21] As can be seen the models tend to estimate these values reasonably well. For instance, for districts with $434 \leq N \leq 900$, observed turnout is 43.2%, the re-estimated CCM model predicts 43.3%, and our model with $m = 1$ predicts 44.5%.[22]

Table 2 instead reports the probability of "fat tails" by electoral district size, comparing this data with the predictions of the pivotal voter model by CCM and our model.[23] Here the underestimation is very significant. For instance, margins of victory above 20% are common in the data, but essentially never occur in the estimated models without groups. For districts with $434 \leq N \leq 900$, the probability of a margin of victory larger than 20% is 12.5%; it is essentially zero in both the CCM model and our model with $m = 1$. Figure 2 in Section 5 highlights this point illustrating the entire empirical distribution of the margin of victory and comparing with the empirical predictions for the full sample.

We can summarize this observation as follows:

[20] A similar finding was obtained in the earlier Hansen et al. (1987) study of voter turnout in Oregon school bond referenda.

[21] We retain all 141 available elections with $0 < N \leq 900$, including the seven with $N < 100$, and use cutoffs of 247 and 434 for reporting. In Section 4, we will show that their analysis is robust to extending the dataset over time and to districts of all sizes. A note on provenance: we use here the sample from Coate and Conlin (2004), which is freely available online. Coate et al. (2008), using the same sample description, contains three more elections for the relevant period, which we could not recover.

[22] The last column of the table also shows the estimated expected turnout with groups. These results will be discussed in Section 3.2.3 and 5.2.

[23] Coate et al. (2008) only report a figure of the distribution of the margin of victory. The specific values reported in Table 2 are computed from the available replication sample and our re-estimated models. A detailed discussion of the Texas Liquor referenda is presented in Section 4.

Table 2: CCM-Sample Tail Probabilities

	Data	CCM	$m = 1$	Grouped
$N < 247$, 45 elections				
$\Pr(\|M\| > .05)$	0.800	0.509	0.618	0.768
$\Pr(\|M\| > .10)$	0.667	0.195	0.300	0.553
$\Pr(\|M\| > .15)$	0.511	0.066	0.115	0.381
$\Pr(\|M\| > .20)$	0.467	0.024	0.039	0.253
$\Pr(\|M\| > .25)$	0.333	0.009	0.013	0.154
$247 \leq N < 434$, 48 elections				
$\Pr(\|M\| > .05)$	0.812	0.278	0.459	0.699
$\Pr(\|M\| > .10)$	0.646	0.026	0.092	0.431
$\Pr(\|M\| > .15)$	0.479	0.001	0.007	0.239
$\Pr(\|M\| > .20)$	0.271	0.000	0.000	0.117
$\Pr(\|M\| > .25)$	0.208	0.000	0.000	0.052
$N \geq 434$, 48 elections				
$\Pr(\|M\| > .05)$	0.771	0.118	0.282	0.626
$\Pr(\|M\| > .10)$	0.438	0.002	0.014	0.321
$\Pr(\|M\| > .15)$	0.208	0.000	0.000	0.135
$\Pr(\|M\| > .20)$	0.125	0.000	0.000	0.048
$\Pr(\|M\| > .25)$	0.083	0.000	0.000	0.015

Notes: All 141 available CCM elections with $0 < N \leq 900$ are retained, including seven with $N < 100$. The fixed bins are $N < 247$, $247 \leq N < 434$, and $N \geq 434$. Margin of victory is $(V_R - V_L)/N$; model tails use 1,000 predictive draws per election.

Observation 1 *The empirical distribution of the Margin of Victory in the Texas liquor referenda studied by CCM has fat tails which do not seem compatible with the predictions of the Pivotal Voter Model.*

Coate et al. (2008) do not have a formal characterization of the distribution of turnout and the margin of victory, so they base their observation on the empirical evidence and on a numerical simulation of their estimated model in a specific dataset. Our model allows us to formalize and generalize Observation 1, thus helping to better understand the source of the problem. When $m = 1$, we have only two variables to estimate: γ_R and γ_L. In the ideal calibration described above, these variables can be uniquely recovered by solving (13) and (15). Because expected turnout is used to calibrate γ_R and γ_L, the model matches that expected turnout by construction. The model predicts that the number of votes cast for option j is a Poisson random variable with mean μ_j. Total turnout is therefore distributed according to a Poisson distribution with mean $\mu = \mu_R + \mu_L$ and standard deviation $\sigma_T = \sqrt{\mu_R + \mu_L}$. The margin of victory is the difference between the vote counts for R and L and is therefore distributed according to a Skellam distribution $S_i(\mu_R, \mu_L)$, with mean $\mu_{MV} = \mu_R - \mu_L$ and standard deviation $\sigma_{MV} = \sqrt{\mu_R + \mu_L}$.[24]

A standard measure of relative dispersion for a distribution with positive mean μ and stan-

[24] In general, given two independent Poisson random variables $\widetilde{x}_1$ and $\widetilde{x}_2$ with mean μ_1 and μ_2, the random variable $\widetilde{x}_1 - \widetilde{x}_2$ is a Skellam distribution with mean $\mu_1 - \mu_2$ and variance $\mu_1 + \mu_2$.

dard deviation σ is the *coefficient of variation*, defined as $CV = \sigma/\mu$ (Lovie (2005)). Using the observations above, we obtain:

Proposition 5. *In the pivotal voter model, we have* $\sigma_T/\mu \to 0$ *and* $\sigma_{MV}/\mu \to 0$ *as* $\mu \to \infty$.

Proof. See Appendix.

Proposition 5 clarifies why it is hard for the pivotal voter model to fit all the key moments of interest. The problem is not the size of the pivotal probability. This is certainly small, but there is no reason to expect ex ante that the net costs are so large to be incompatible with this. The cost of voting or its distribution is an unknown variable to be estimated. The estimation presented by CCM and the estimation we will discuss at length in Section 4 show that an estimated pivotal voter model does not find it difficult to fit expected turnout. The real problem is that if we calibrate a model to fit turnout, then the remaining degrees of freedom in the distribution are minimal. This happens as a consequence of the assumption the agents vote as *independent units*. Once the probability a voter or group votes is pinned down (i.e., θ_j for $j = L, R$), then expected turnout is also pinned down (i.e., $\theta_j N$ for $j = L, R$), and its standard deviation (i.e., $\sqrt{\theta_j N}$ for $j = L, R$): and thus the shape of the distribution of the margin of victory.

It is important to stress that this is a fundamental property of voting models, not a feature of the specific model we are using (and indeed this explains why both our model and the significantly different model adopted by Coate et al. (2008) share the same property).

If, instead of assuming a Poisson population, we assume a fixed population of N voters, the same concentration result holds whenever individual voting decisions are independent conditional on election characteristics. Evaluating all moments conditional on these characteristics and writing $\mu = \mu_R + \mu_L$, we have:

$$\sigma_T^2 \leq \mu, \qquad \sigma_{MV}^2 \leq \mu.$$

Consequently,

$$\frac{\sigma_T}{\mu} \leq \frac{1}{\sqrt{\mu}}, \qquad \frac{\sigma_{MV}}{\mu} \leq \frac{1}{\sqrt{\mu}},$$

and both ratios converge to zero as expected turnout increases. Thus, as in the Poisson model, independent voting limits dispersion around the conditional expected margin relative to expected turnout.[25] Whether a model can account for the observed large margins also depends on its predictions for that expected margin. The results, therefore, would be very similar.[26] In both

[25] Let $Z_i \in \{-1, 0, 1\}$ denote voter i's contribution to the signed vote-count margin, and let $A_i = |Z_i|$ indicate participation. Conditional independence across voters and the identities $Z_i^2 = A_i^2 = A_i$ imply $\mathrm{Var}\left(\sum_i Z_i\right) \leq \sum_i \mathbb{E}[A_i] = \mu$, and $\mathrm{Var}\left(\sum_i A_i\right) \leq \sum_i \mathbb{E}[A_i] = \mu$.

[26] Indeed, Coate et al. (2008) highlight the discrepancy between the theoretical model with a constant N and the data.

Table 3: National Two-Candidate Elections, 1980–2024

Statistic	All offices	US House	US Senate
Elections	5,008	4,703	305
Median eligible voters	457,136	446,018	2,980,672
Turnout	0.495	0.495	0.508
Mean abs. margin	0.137	0.140	0.098
$\Pr(\lvert M\rvert > 0.05)$	0.818	0.829	0.659
$\Pr(\lvert M\rvert > 0.10)$	0.617	0.630	0.420
$\Pr(\lvert M\rvert > 0.15)$	0.405	0.416	0.226
$\Pr(\lvert M\rvert > 0.20)$	0.225	0.232	0.111
$\Pr(\lvert M\rvert > 0.25)$	0.109	0.114	0.043

Notes: The signed margin is $M = (V_R - V_L)/N$ and N is eligible voters. Races are restricted to elections with exactly two candidates receiving at least 1 percent of the election vote total.

models, the standard deviation of turnout, σ_j, is approximately equal to $\sqrt{\mu_j}$, the square root of the expected turnout for j: implying that $\sigma_j/\mu_j \to 0$ as $\mu_j \to \infty$. This implies that in elections with even moderate turnout, the ratio is small, smaller than empirically observed.

Two additional remarks highlight the generality of the issues discussed above. First, the difficulties in rationalizing the data are not an exclusive feature of pivotal-voter models with $m = 1$; rather, they affect most models based on the assumption that voters act independently. To see this, consider a different theory of elections that predicts that voters for j in district k vote with some probability $\beta_j(\gamma, X_k)$, which may depend on unobserved preferences and other observable characteristics of the district. For instance, the theory that determines $\beta_j(\gamma, X_k)$ could be Ferejohn and Fiorina's *Regret Maximizer Model* or Feddersen and Sandroni's *Ethical Voter Model*. If the estimated model fits the observed expected vote count $\widetilde{\mu}_{j,k}$, then $\beta_j(\gamma, X_k) = \widetilde{\mu}_{j,k}/N_k$. Then, as an implication of independent choice behavior, the vote counts for R and L are Poisson random variables (or binomial random variables in a fixed-population formulation) with respective means $\widetilde{\mu}_{R,k}$ and $\widetilde{\mu}_{L,k}$. The Margin of Victory is therefore the difference between these two vote counts. In the Poisson case, it has mean $\mu_{MV,k} = \widetilde{\mu}_{R,k} - \widetilde{\mu}_{L,k}$ and standard deviation $\sigma_{MV,k} = \sqrt{\widetilde{\mu}_{R,k} + \widetilde{\mu}_{L,k}}$; in the binomial case, the same variance expression is the usual approximation when the individual voting probabilities are small. Hence, defining $\widetilde{\mu}_k = \widetilde{\mu}_{R,k} + \widetilde{\mu}_{L,k}$, we obtain $\sigma_{MV,k}/\widetilde{\mu}_k \simeq 1/\sqrt{\widetilde{\mu}_k}$, which is approximately zero for large $\widetilde{\mu}_k$.[27]

[27] A degenerate distribution of MV is also predicted in Pogorelskiy (2024)'s model of high-turnout correlated equilibria, although the logic is slightly different. The equilibria constructed in Pogorelskiy (2024) involve strategies that generate nearly identical turnout levels across the two sides and therefore cannot explain the large observed margins of victory discussed above. Still, the emphasis on coordinated behavior in that work is related to our approach. The main difference is that coordination in our model occurs within many small social groups, rather than through correlated equilibria involving two electorate-wide groups, one for each side.

Second, the presence of "fat tails" in the distribution of the margin of victory is not just a problem with Texas liquor referenda. Table 3 illustrates the distribution of the margin of victory for Congressional and Senatorial races. In this data too, values of the MV larger than 20% are a common occurrence even in large elections.[28] Even without a full-fledged estimation of the model, Proposition 5 provides a useful benchmark for interpreting the congressional and senatorial data. A model with independent voters predicts that the standard deviation of the margin of victory grows only with the square root of expected turnout, so that σ_{MV}/μ becomes small as turnout increases. Table 4 compares this benchmark with the dispersion observed across elections within electorate-size bins. Because within-bin dispersion may also capture some heterogeneity in expected electoral margins across elections, the comparison is best viewed as an informative empirical benchmark rather than a literal test of Proposition 5. Nevertheless, the discrepancy is striking: the benchmark values of σ_{MV}/μ are 0.2% and 0.09% for U.S. House and Senate races, respectively, whereas the corresponding values in the data are 33.4% and 22.36%.

We can extend Observation 1 in light of Proposition 5 and this evidence as follows:

Observation 2 *The empirical distributions of the Margin of Victory in two-candidate elections in the U.S. and liquor referenda in Texas exhibit much greater dispersion and substantially fatter tails than the homogeneous independent-voter benchmark characterized by Proposition 5.*

In the next section we discuss why groups may help reconcile this general discrepancy between theory and data. Before we turn to the empirical analysis, we close this section with a final remark. In Coate et al. (2008), as in our model and in most other standard pivotal-voter models, there are no aggregate shocks. The models, however, can be easily extended to include aggregate shocks to voting costs, the distribution of information, or voters' preferences.[29] Before turning to the importance of groups, it is worth discussing why allowing for aggregate shocks is unlikely to provide a satisfactory explanation of the evidence under plausible assumptions.

Aggregate shocks may be of two types: they may be observable, as when a change in legislation changes the cost of voting, or a scandal makes a candidate unpopular; or unobservable, when a shock changes the share of supporters for a candidate that is undetected even by polls. When the shock is observed before voting, the analysis changes little: voters condition their behavior on its realization. For example, if there is a shock to the distribution of preference and a larger share of right-wing voters is expected, right-wing voters would endogenously participate with lower

[28] The national sample contains 4,703 U.S. House elections and 305 U.S. Senate elections. We treat a candidate as major if the candidate receives at least 1 percent of the election vote total, and keep elections with exactly two major candidates.

[29] Hansen et al. (1987) consider a model with observed aggregate shocks to the distribution of voting costs; Battaglini et al. (2010) consider a model with unobserved uncertainty about the distribution of voters' information; and Coate and Conlin (2004) introduce unobserved shocks to the distribution of voters' political affiliations.

Table 4: National Election Moment Scaling by Two-Party Vote Totals

Sample	Elections	Bins	Median σ/μ	Median $1/\sqrt{\mu}$	$\Pr(\lvert M\rvert > 0.05)$	$\Pr(\lvert M\rvert > 0.10)$	$\Pr(\lvert M\rvert > 0.20)$
All offices	5,008	20	0.3298	0.0021	0.818	0.617	0.225
US House	4,703	20	0.3342	0.0022	0.829	0.630	0.232
US Senate	305	20	0.2236	0.0009	0.659	0.420	0.111

Notes: For each sample separately, elections are sorted by realized two-party votes and divided into 20 approximately equal-count bins. Within each bin, μ is mean two-party votes and σ is the within-bin standard deviation of the signed two-party margin, $V_R - V_L$. The two median columns give unweighted medians across the 20 bins. The independent-voter benchmark has $\sigma/\mu = 1/\sqrt{\mu}$. The last three columns report the empirical probabilities that the absolute eligible-voter margin $|M|$ exceeds the indicated thresholds, as in Table 3.

probability and left-wing voters with higher probability: the margin of victory would not increase in equilibrium.[30] This is the "underdog effect" emphasized by Levine and Palfrey (2007). Unobserved aggregate shocks, however, may help rationalize large margins of victory in elections with high turnout. Following Feddersen and Pesendorfer (1996), Battaglini et al. (2010) study a model in which the optimal ballot depends on the state, some voters are uninformed, and a random number are perfectly informed. Uninformed voters choose their strategies without observing the number of informed voters and participate only to offset an ex ante bias toward the opposing side, such as that created by partisans who always vote for one option. Their votes for the two options therefore balance in expectation, while the realized margin is determined by the number of informed voters. The larger this realization, the larger the margin of victory. Similarly, Coate and Conlin (2004) consider a two-candidate election, with candidates R and L, in which not only the realized number of R-supporters, its expected value is random as well. Voters choose participation without observing the realized mean of the preference distribution and therefore cannot adjust to it. High realizations of R-supporters generate large victories for R, and analogously for L.

As we discuss in Section 5.2, this mechanism may help explain unusually large margins of victory, but is unlikely to fully account for the large margins observed in the data. Although it is natural for, say, the realized number of informed voters or R-supporters to be random, it is less plausible that there is large uncertainty on the expected value, that is there are shocks that have a large impact on the distribution of preferences (or costs of information), but remain unobserved. Within a district, voters are likely to have reasonably accurate information from past elections, polls, and other public signals. Put simply, voters know whether they are voting in Texas or California. Indeed, in Section 5.2.2 we will explicitly compare a version of Coate and Conlin (2004) with unobserved shock to the distribution of preferences and independent voters to our model with

[30] Indeed, it can be seen that in our model the shares α_R and α_L do not play an important role in (9)-(10), which pin down $\frac{\alpha_R}{m} N\beta_R$ and $\frac{\alpha_L}{m} N\beta_L$ irrespective of the specific values α_R and α_L. A higher value of α_R would only imply a lower β_R.

no aggregate shock and groups, showing the latter performs better.

3.2.3 Why organized groups help explain the evidence

The picture changes if there are organized groups of voters. In this case, under the assumptions of Propositions 2 or 4, only the organized groups vote and the number of groups that vote for j, N_j, is a Poisson with mean and variance equal to $(\theta_j N)/m_N$, where m_N is the size of the groups in a district of size N. Expected turnout for candidate j is therefore $\mu_j(N) = m_N(\theta_j N/m_N) = \theta_j N$. Similarly to the classical case of $m_N = 1$, total expected turnout is:

$$\mu_N = \mu_{N,L} + \mu_{N,R} = \theta_L N + \theta_R N \tag{16}$$

which is independent of m_N holding θ_L and θ_R fixed. The variance of turnout, however, is:

$$\sigma_N^2 = m_N^2 \left[\frac{\theta_L N}{m_N} + \frac{\theta_R N}{m_N}\right] = m_N \left[\mu_{NR} + \mu_{N,L}\right] \tag{17}$$

which is now linear in the size of the groups m_N. We do not observe m_N, but we can recover it from the data using (14) if we observe the empirical variance of the margin of victory $\widetilde{\sigma}_N^2$ and the empirical expected value $\widetilde{\mu}_N$.

As before, the difference in the numbers of voting groups, $N_R - N_L$, follows a Skellam distribution. The vote margin $m_N(N_R - N_L)$ is therefore a scaled Skellam random variable, with mean $\mu_{N,R} - \mu_{N,L}$ and standard deviation $\sqrt{m_N(\mu_{N,R} + \mu_{N,L})}$, which can be estimated as

$$\begin{aligned} \mu_{N,MV}(m_N) &= \mu_{N,R} - \mu_{N,L} \\ \text{and } \sigma_{N,MV} &= \left[m_N^2\left(\frac{\theta_L N}{m_N} + \frac{\theta_R N}{m_N}\right)\right]^{1/2} = \sqrt{m_N(\mu_{N,R} + \mu_{N,L})} \end{aligned}$$

The net result of introducing groups is that, holding expected turnout fixed, the variances of turnout and of the margin of victory are increased approximately by a factor of m_N relative to the case without organized groups. This helps rationalize the data since, depending on the size of groups, it may help explain the excess volatility and fat tails in the margin of victory. On the other hand, the group size m_N is not simply a free variable in the model, since it is pinned down by (14). The ratio $\sigma_{N,MV}/\mu_N$ is now $\sqrt{m_N/(\mu_{N,R} + \mu_{N,L})}$: as N increases, $\widetilde{\mu}_R + \widetilde{\mu}_L$ typically increases too; but m_N also increases, offsetting or at least attenuating the effect on the ratio. In the next section, we estimate the model and study the extent to which allowing for groups improves its fit.

4 Estimating the model

4.1 Texas liquor referenda

We estimate our model using data from Texas liquor referenda. As mentioned above, these data have been used extensively to study participation in elections because they provide a dataset of

uniform elections, with widely varying city sizes, all concerning essentially a single issue (see, for instance, Coate and Conlin (2004), Coate et al. (2008), and Arnold and Sieg (2026)). As an additional benefit, using the same data as previous work makes the assessment of our results easier to interpret.

Texas regulates alcohol sales through a statewide licensing system overlaid on a local-option regime. Under Article XVI, Section 20 of the Texas Constitution, voters in a county, municipality, or justice precinct (a subdivision of counties) may determine which categories of alcoholic beverages may be sold within that territory directly. The framework has been in place in essentially modern form since the end of prohibition. As a consequence, Texas does not have a single county-level wet/dry map. Alcohol status is territorially fine-grained, and different parts of the same county can face different rules.

We refer to Coate and Conlin (2004) for a complete discussion of the institutional details of these referenda. The common policy issue makes them useful for studying turnout, although other contests may appear on the same ballot. As a robustness check, we exclude elections held on standard November and regular primary election dates. The evidence for grouped voting persists in this restricted sample (Appendix II). This calendar restriction is a proxy for concurrent elections; it does not establish which other contests appeared on each ballot.

Following Coate and Conlin (2004) and Coate et al. (2008), we focus the analysis on referenda in which the jurisdiction was fully "dry" (prohibiting the sale of all alcoholic beverages) before the referendum, meaning that a vote "for" (denoted L) is always liberalizing and a vote "against" (denoted R) is always maintaining the *status quo*.[31] Considering data from 1937 to 2019, this gives us a total of 1,724 referenda, comprising districts from 102 to 156,796.5 estimated eligible voters.[32]

Election-level data—including the jurisdiction name and county, election date, vote totals for and against legalization, the type of legalization proposed, and whether the jurisdiction is a city or a county/justice precinct—were extracted from the Texas Alcoholic Beverage Commission (TABC) annual reports, which are available as scanned PDFs covering the period 1937–2020. Figure 7 in the Appendix shows the distribution of the number of dry-to-wet referenda over the years.

Because the TABC reports do not record the number of registered or eligible voters, we construct the eligible-population estimate N_i for district i from external demographic sources. In empirical size cutoffs and covariates, N denotes this same quantity with the district index suppressed. For

[31] Liquor referenda can also move in the opposite direction, from wet to dry, but we focus on dry-to-wet referenda for consistency with the previous literature. Because Texas liquor referenda may concern different categories of alcohol sales, we control for the type of ballot question in the dataset.

[32] Coate and Conlin (2004) restricted attention to 366 referenda from 1976 to 1996; Coate et al. (2008) further restricted attention to 144 referenda involving districts with fewer than 900 citizens. The time window of our dataset is determined by the availability of the underlying records: the Texas Alcoholic Beverage Commission annual reports begin in 1937.

county and justice precinct[33] elections, we combine age-by-county tabulations from NHGIS decennial censuses (1940, 1950, 1960) with SEER county population estimates by age (1969–2019), computing the voting-age population as the population aged 20 and above plus two-fifths of the population aged 15–19 from July 1, 1971 onward, and the population aged 25 and above plus four-fifths of the population aged 20–24 through June 30, 1971 (the voting age was 21 until the 26th amendment's ratification). We interpolate linearly between census years and extrapolate beyond the observed range using constant values. For city elections, we use the NHGIS place-level time series (1970–2020 decennial censuses) for the population aged 18 and above, supplemented by Texas Almanac city population data (1850–2000) scaled by the city-specific voting-age share.[34]

An important variable for determining voters' preferences and assessing the presence of relevant social structures in local communities is voters' religious affiliation. To measure this, we construct the percentage of the county population that is Southern Baptist using the Longitudinal Religious Congregations and Membership File, 1980–2010, which reports county-level adherent counts by denomination. We compute the ratio of Southern Baptist Convention adherents to total county population, interpolate linearly between the decadal observations, and extrapolate using the nearest available value for years outside the 1980–2010 window. We additionally use standardized Southern Baptist average congregation size from the same source. We exclude elections with missing values in any of the key variables (eligible voters, vote totals, Baptist population share, or Baptist congregation size) and elections with empirical turnout exceeding one.[35] Our baseline analysis uses the sample with electorate size $N \geq 100$.

To investigate heterogeneity in the model's parameters across elections of different scales, we partition the sample into three strata by electorate size: small, $N < 1{,}500$; medium, $1{,}500 \leq N < 5{,}000$; and large, $N \geq 5{,}000$. These cutoffs are rounded electorate cutoffs near terciles, so each stratum contains roughly one third of the estimation sample. We estimate the models both on the full sample and separately within each stratum. Table 5 reports summary statistics for the main structural estimation sample. Figure 1 plots turnout as a fraction of eligible voters against electorate size (on a log scale).

The specifications assuming individual voting ($m = 1$) use a parsimonious set of covariates that are available for the full historical sample and match those used by Coate et al. (2008): an intercept, $\log N$, the county Southern Baptist population share, an indicator for city elections,

[33] A justice precinct is a uniform division of a county into a fixed number of precincts. For these elections, we divide the county population by the total number of justice precincts in the relevant year.

[34] A caveat is that no reasonable citizenship records exist for the time period, so our estimates of the eligible voter population are necessarily *over*estimates because we cannot exclude noncitizens.

[35] A relic of the nature of estimating the number of eligible voters in the population.

Table 5: Summary Statistics: Main Structural Estimation Sample

	Full ($E = 1{,}724$)	Small ($E = 640$)	Medium ($E = 561$)	Large ($E = 523$)
Estimated electorate denominator (N)	7,174 (15,136)	698 (386)	2,853 (982)	19,735 (22,923)
Votes for	716 (1,650)	119 (85)	345 (266)	1,843 (2,655)
Votes against	754 (1,751)	125 (100)	374 (304)	1,930 (2,826)
Turnout rate	0.292 (0.181)	0.387 (0.187)	0.255 (0.159)	0.214 (0.141)
Margin of victory	−0.009 (0.102)	−0.005 (0.137)	−0.009 (0.078)	−0.013 (0.071)
Pct. Baptist	0.276 (0.109)	0.282 (0.107)	0.280 (0.116)	0.265 (0.104)
Avg. SBC congregation size	526 (267)	457 (217)	488 (231)	652 (311)
City election	0.554 (0.497)	0.792 (0.406)	0.446 (0.497)	0.379 (0.485)
Weekend election	0.645 (0.479)	0.662 (0.473)	0.643 (0.479)	0.625 (0.485)
All-alcohol/non-mixed	0.266 (0.442)	0.353 (0.478)	0.275 (0.447)	0.149 (0.357)
Mixed-beverage unrestricted	0.191 (0.393)	0.203 (0.403)	0.193 (0.395)	0.174 (0.379)
Restaurant-only mixed-beverage	0.085 (0.278)	0.031 (0.174)	0.094 (0.293)	0.140 (0.347)
For-side wins	0.469 (0.499)	0.495 (0.500)	0.446 (0.497)	0.461 (0.499)

Notes: Cells report means with standard deviations in parentheses; binary variables are reported as proportions. Baseline sample: previously dry local-option elections that pass the complete-case and turnout screens and have estimated eligible voters of at least 100. Estimated eligible voters min/median/max: Full: 102 / 2,370 / 156,796; Small: 102 / 651 / 1,494; Medium: 1,502 / 2,659 / 4,988; Large: 5,014 / 10,996 / 156,796. Small, medium, and large jurisdictions are defined by the fixed size cutoffs: $N < 1500$, $1500 \leq N < 5000$, and $N \geq 5000$.

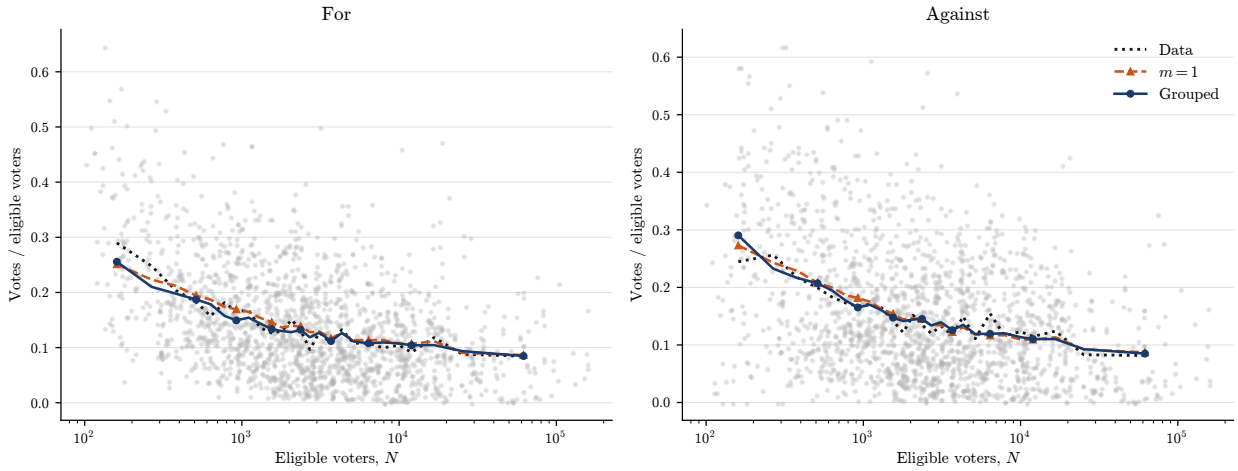


Figure 1: Turnout by electorate size in the Texas main sample. The two panels report votes for (left) and against (right) as a fraction of eligible voters; in each panel the empirical pattern is compared with fitted turnout from the $m = 1$ benchmark and the grouped model.

an indicator for weekend elections, and indicators for the broad alcohol proposition types.[36] The grouped model uses the same cost/benefit ratio covariates, and adds a separate group-size equation with an intercept, $\log N$, $[\log(N)]^2$, city status, Baptist share, and standardized average Southern Baptist congregation size. Appendix II reports estimates by period, restrictions to first referenda, alternative electorate-size thresholds, and richer covariates. The grouped model fits better than

[36]Coate et al. (2008) use only two indicators, one for all alcohol on-premises and one for on- and off-premeses consumption of all alcohol, excluding beer and wine. Our analysis of the referendum types exposed an additional category, so we use four mutually-exclusive categories: beer and wine only, all alcohol on- or off-premises, mixed beverages on- or off-premises, and mixed beverages only in restaurants, excluding the beer and wine only group.

the $m = 1$ benchmark throughout these checks, although the estimated magnitude of group size varies across specifications.

4.2 Structural Estimation

We estimate the cost/value ratios $\gamma_{L,i}$ and $\gamma_{R,i}$ and group size m_i as functions of election characteristics. Using the covariates described in Section 4.1, we specify

$$\log \gamma_{R,i} = X_i^\top \xi_R, \qquad \log \gamma_{L,i} = X_i^\top \xi_L, \qquad \log(m_i/N_i) = X_{m,i}^\top \xi_m.$$

The two cost/value equations have separate coefficient vectors. The group-share equation applies before rounding; we bound and round its implied group size to an integer between 1 and $\lfloor N_i \rfloor$.

Let $Y_{s,i} \sim \text{Poisson}(\lambda_{s,i})$ denote the number of participating groups on side s. The equilibrium conditions (1)–(2) determine their intensities directly from the cost/value ratios:

$$\gamma_{s,i} = \frac{\exp\left[-(\sqrt{\lambda_{R,i}} - \sqrt{\lambda_{L,i}})^2\right]}{4\sqrt{\pi}(\lambda_{R,i}\lambda_{L,i})^{1/4}} \left(1 + \sqrt{\frac{\lambda_{-s,i}}{\lambda_{s,i}}}\right), \qquad s \in \{R, L\}.$$

Here $-s$ denotes the other side, and expected strategic vote shares are $\theta_{s,i} = m_i\lambda_{s,i}/N_i$. In the grouped equilibrium, $\theta_{s,i} = \alpha_{s,i}\beta_{s,i}$. We restrict $0 < \gamma_{s,i} < 1/2$ and $0 < m_i(\lambda_{R,i} + \lambda_{L,i}) < N_i$, so expected turnout is interior.

In the empirical model, observed vote counts satisfy

$$T_{s,i} = m_i Y_{s,i} + \varepsilon_{s,i}, \tag{18}$$

where $\varepsilon_{s,i}$ is a mean-zero noise term constructed as a simmetric Skellam distribution with mean zero and variance of 1% of N_i.[37] The advantage of using a symmetric Skellam distribution is that it yields an integer-valued, bell-shaped, mean-zero noise term that is approximately normally distributed for large N_i. The noise is part of the statistical observation equation, not the underlying voting game. It provides an approximation to the distribution of observed counts around the model's latent block-vote totals and does not represent additional citizens casting ballots. Appendix II.5 discusses the noise term and examines alternative distributions (such as a Poisson distribution shifted to have zero mean) and alternative component means, including one that does not scale with electorate size. All results are qualitatively similar.

Writing $\xi = (\xi_R, \xi_L, \xi_m)$ and $\widetilde{\theta}_{s,i} = V_{s,i}/N_i$ for the observed vote share, we maximize

$$\ell(\xi) = \sum_i \sum_{s \in \{R,L\}} \log \Pr_{\xi}\Big(m_i Y_{s,i} + \varepsilon_{s,i} = N_i \widetilde{\theta}_{s,i}\Big).$$

[37] The Skellam noise term is generated as the difference between two independent Poisson random variables, each with mean $0.005 \cdot N_i$.

The probability sums over the possible numbers of participating groups. We compare the restriction $m_i = 1$ with a specification allowing group size to vary. The noise adds $0.02N_i$ to the variance of both total votes and the vote-count margin, without changing their means. Under the main rule, the probability that absolute net noise across both sides exceeds 5% of the electorate is below 0.11% in every election in the main sample.

We use 200 county-block bootstrap draws for inference, holding the noise rule fixed. Appendix III.1 provides the likelihood, optimization, and inference procedures.

5 Empirical Results

In this section, we present the results of the structural estimation of our model. Section 5.1 considers the model without organized groups, that is, the restricted specification with $m = 1$. We compare its estimates with those obtained by Coate, Conlin, and Moro (CCM), using both their original data set (1976-1996) and our extended one (1937-2019). This comparison provides an important benchmark for evaluating our estimates and validating our estimation procedure, as the two approaches yield very similar qualitative results. In particular, our analysis of the restricted model confirms the two main findings of CCM: the estimated model predicts expected turnout well, both in and out of sample, but it cannot account for the observed magnitude and distribution of the margin of victory.

In Section 5.2, we estimate the full model with $m \geq 1$, allowing for organized groups. We emphasize three sets of results. First, groups substantially improve the fit of turnout and the margin-of-victory distribution, although the model still understates the observed tails. The no-groups model is a restriction of the grouped model, and the likelihood-ratio statistic is $\ell = 916{,}474.26$, implying rejection of the null at every conventional significance level. A fitted-null simulation reference places this gain far beyond the variation generated by the estimated no-groups model.[38] Second, we examine what the estimates reveal about group size and bounds on the organized share of the electorate. Third, we compare our model with alternative, non-nested voting models estimated on the same data. Sections 5.3 and 5.4 summarize the robustness and out-of-sample exercises reported in Appendix II.

[38] Integer rounding and the parameters associated with grouping preclude an automatic chi-square calibration. We simulate 99,999 datasets from the fitted $m = 1$ model, holding covariates and the noise rule fixed. A conservative upper bound on each simulated likelihood-ratio statistic is below the observed statistic, giving an add-one upper-tail estimate no greater than 10^{-5}.

Table 6: Turnout and Tail Probabilities by Size Group. The small, medium, and large strata are defined by $N < 1{,}500$, $1{,}500 \leq N < 5{,}000$, and $N \geq 5{,}000$ eligible voters.

Stratum	Series	Turnout	$\Pr(\lvert M\rvert > .05)$	$\Pr(\lvert M\rvert > .10)$	$\Pr(\lvert M\rvert > .15)$	$\Pr(\lvert M\rvert > .20)$
Full	Empirical	0.292	0.485	0.247	0.128	0.068
	CCM	0.402	0.105	0.036	0.012	0.004
	$m = 1$	0.299	0.049	0.006	0.001	0.000
	Grouped	0.290	0.401	0.141	0.052	0.020
Small	Empirical	0.387	0.656	0.386	0.219	0.134
	CCM	0.672	0.281	0.096	0.033	0.010
	$m = 1$	0.389	0.128	0.021	0.003	0.000
	Grouped	0.386	0.566	0.266	0.114	0.049
Medium	Empirical	0.255	0.408	0.176	0.078	0.034
	CCM	0.320	0.003	0.000	0.000	0.000
	$m = 1$	0.260	0.002	0.000	0.000	0.000
	Grouped	0.256	0.390	0.095	0.016	0.002
Large	Empirical	0.214	0.359	0.153	0.069	0.023
	CCM	0.161	0.000	0.000	0.000	0.000
	$m = 1$	0.224	0.000	0.000	0.000	0.000
	Grouped	0.213	0.238	0.028	0.002	0.000

Notes: 1,724 elections (640 small, 561 medium, 523 large). Model moments are computed from simulated elections using the main fitted specifications. CCM uses the same fitted predictive object as Table 8, restricted to each stratum. Margin of victory is the observed margin in the data and is simulated 1,000 times in the simulations.

5.1 Elections with no organized groups

Both our restricted model without groups, that is, with $m = 1$, and the model in Coate et al. (2008) are variants of the classic pivotal-voter model, in which voting is purely instrumental. They differ, however, in several important respects. Coate et al. (2008) assume privately observed voting costs and a binomial distribution for the number of supporters, as in Palfrey and Rosenthal (1985). We instead assume publicly observed voting costs and Poisson strategic arrivals, as in Myerson (1998), with the mean-zero observation noise specified above. Despite these differences, both approaches capture the decline in turnout with electorate size while substantially understating large margins of victory.

Tables 1 and 2, that were briefly discussed in Section 3.2.2, compare the re-estimated CCM benchmark with our model on the 141 available replication elections with $0 < N \leq 900$. We return to them here to highlight their similar performance: both models do a good job estimating expected turnout in the CCM sample; but they fail to account for the empirical distribution of the margin of victory. For instance, in the largest bin, $434 \leq N \leq 900$, a margin of victory larger than 15% has an empirical probability of 20.8%: the CCM model, however, predicts essentially a zero probability for this event; and our model with $m = 1$, predicts a probability below 0.1%.

For a better comparison, we evaluate the CCM model using our full data set, which was not available and was beyond the computational capacities of computers when CCM was written. Estimates of turnout by district size for the full sample are reported in Table 6. The table again shows

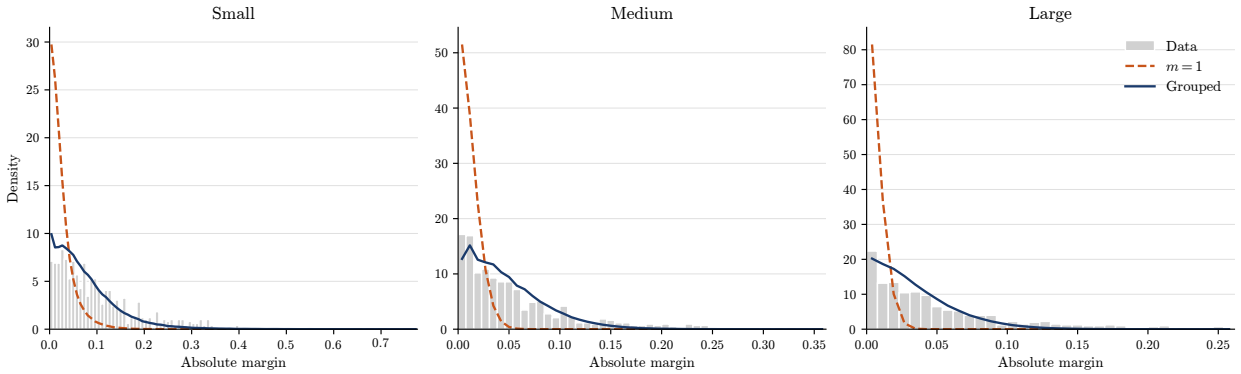


Figure 2: Absolute margin of victory by electorate-size stratum. Panels report, from left to right, small, medium, and large jurisdictions; in each panel the empirical distribution is compared with simulations from the fitted $m = 1$ benchmark and, anticipating the results of the next section, from the fitted grouped-vote model.

that both pivotal-voter models capture the sharp decline in turnout as electorate size increases. Average turnout in the data falls from 38.7% in small jurisdictions to 25.5% in medium jurisdictions and 21.4% in large jurisdictions; both models reproduce this downward gradient: the CCM model predicts turnout rates of 67.2%, 32.0%, and 16.1%, respectively, while our restricted model with $m = 1$ predicts 38.9%, 26.0%, and 22.4%. Thus, although the CCM model overpredicts turnout in small jurisdictions, both models track the decline in turnout across jurisdiction sizes, and our restricted model also matches the observed levels closely.

Table 6 also presents the tail results for the margin of victory. The table shows that the same problem found in the CCM data is also present in the full sample, if anything in even more extreme form. Margins of victory above 5% are common even outside the small-election stratum: in medium-size jurisdictions ($1{,}500 \leq N < 5{,}000$) they occur with probability 40.8% in the data, while the $m = 1$ benchmark predicts 0.2% and CCM predicts 0.3%. In large jurisdictions ($N \geq 5{,}000$), these margins occur with probability 35.9% in the data, while the $m = 1$ benchmark predicts below 0.1% and CCM again predicts zero. Even among small elections ($N < 1{,}500$), margins of victory above 15% occur with probability 21.9% in the data. By contrast, they occur with probability 0.3% in the $m = 1$ benchmark and 3.3% in CCM. Figure 2 illustrates this phenomenon reporting the distribution of the absolute value of the margin of victory and compares it with the estimated models using the full empirical sample, conditioning on three classes of elections: small elections, medium-sized elections, and large elections. Note that while the peaks of the absolute margin of victory distribution are well-predicted, the tails of the simulated model deteriorate to essentially zero much faster than the empirical results. That tail deterioration is exactly the issue identified by Coate et al. (2008).

Table 7: Main Grouped-Vote Estimates

	$\log \gamma_R$	$\log \gamma_L$	$\log(m/N)$
Intercept	-1.112^{***}	-0.925^{***}	0.438
	(0.147)	(0.127)	(0.323)
$\log(N)$	-0.165^{***}	-0.176^{***}	-0.769^{***}
	(0.015)	(0.014)	(0.085)
$[\log(N)]^2$			0.019^{***}
			(0.006)
Baptist share	-0.326^{*}	-0.431^{***}	0.278^{**}
	(0.181)	(0.143)	(0.124)
City	-0.159^{**}	-0.246^{***}	-0.320^{***}
	(0.065)	(0.043)	(0.023)
Weekend	0.122^{***}	0.097^{***}	
	(0.044)	(0.033)	
All-alcohol/non-mixed	-0.252^{***}	-0.144^{**}	
	(0.093)	(0.065)	
Mixed-beverage unrestricted	-0.080	-0.021	
	(0.056)	(0.039)	
Restaurant-only mixed-beverage	-0.049	-0.120	
	(0.135)	(0.180)	
SBC congregation size (std.)			-0.006
			(0.021)
Elections		1,724	
Log likelihood		−29,169.6	
Mean $\hat{m}$		42.26	
Median $\hat{m}$		28.00	

Notes: The native equations are $\log \gamma_s = X^{\top}\xi_s$ for $s \in \{R, L\}$. Before rounding, $\log(m/N) = X_m^{\top}\xi_m$. County-block bootstrap standard errors (200 paired draws) are in parentheses. Coefficient stars use normal approximations: *, **, and *** indicate the 10, 5, and 1 percent levels.

5.2 Elections with organized groups

5.2.1 Fit of the estimate with groups

Table 7 presents the parameter estimates of the full model with organized groups, together with county-block bootstrap standard errors that re-estimate the full model on each resampled set of county clusters. In the full model, in which we allow for groups, the mean estimated group size is about 42.3, with a median of 28.

Allowing for groups raises the full-sample log likelihood from −487,407 under $m = 1$ to −29,170 under the grouped model. Additional parameters necessarily improve the maximized fit of a nested model, but the size of this gain is substantial: the relevant likelihood-ratio statistic is 916,474.26, which rejects the null at all conventional significance levels even accounting for the additional parameters. The fitted-null simulation reference described above provides additional evidence against the no-groups restriction. We compare additional voting models in Section 5.2.3.

The coefficients also describe how voting incentives and group size vary with election characteristics (Table 7; Appendix Table 9 reports separate size-stratum fits). The mean elasticity of turnout with respect to electorate size is −0.15. The Baptist-share coefficients lower the fitted cost/value ratios on both sides. The right side coefficient is significant at ten percent, the left side significant

at one percent. The paired bootstrap contrast between the two coefficients has a nominal p-value of 0.304. These estimates therefore do not establish a cross-side preference difference or separate cost effects from value effects. At the sample mean, increasing Baptist share by one standard deviation, approximately 0.11, raises predicted turnout by about four percentage points. Weekend elections are associated with lower fitted turnout.

Group size increases less than proportionally with the electorate. Before rounding, its fitted elasticity is approximately 0.5 at the median electorate and 0.7 at the largest electorate. The group-share coefficient on Baptist share is positive and significant at the 5% level. The coefficient on average Southern Baptist congregation size is slightly negative, but imprecisely estimated. These associations are consistent with a role for local social organizations, while providing limited evidence on their specific institutional form.

The third observation is that, just as predicted theoretically in Section 3.2.3, the model with groups continues to estimate turnout well. The comparison of empirical turnout with the model predictions for the full sample and by district-size class is presented in the third column of Table 6. This is illustrated in Figure 1, which shows that the full model with groups is able to capture the observed comparative statics of turnout with respect to the size of the electorate.

The fourth observation explains why the model with groups generates a superior fit and confirms the theoretical conjecture presented in Section 3.2.3 that groups can help explain the observed margins of victory. This can be observed from Columns 4-7 of Table 6, that confront the models with the more demanding challenge of matching the distribution of the margin of victory, at thresholds ranging from 5 to 20 percent. As the table shows, the grouped model fits the tails much more accurately than either pivotal-voter benchmark. In the full sample, the probability of a margin of victory larger than 5 percent is 48.5% in the data: it is only 10.5% in the CCM model and 4.9% in our model with $m = 1$, but 40.1% in our model with groups. Figure 2 illustrates this point by showing the distribution of the absolute value of the margin of victory with groups and comparing it with the prediction of the model with no groups.

The ability of the model to fit the data varies with the size of the district. The model performs best for small districts, defined as those with less than 1,500 eligible voters. It also performs relatively well for medium districts, with eligible voters between 1,500 and 5,000: in the data, margins of victory larger than 10% occur with probability equal to 17.6%; in the model with $m = 1$, the probability is below 0.1%; in the model with estimated m, the probability is 9.5%. The performance in large elections is less satisfactory, though still orders of magnitude better than in the $m = 1$ model. Margins of victory larger than 10% occur with probability 15.3% in these elections. The $m = 1$ model predicts essentially zero probability for these events, whereas the

grouped model predicts a probability of 2.8%. Figure 2 displays, separately for small, medium, and large jurisdictions, the empirical distribution of the absolute margin of victory and compares it with the predicted distribution for both the group model and the $m = 1$ model.

The model with groups studied here abstracts from the potential presence of aggregate shocks to, for instance, values or voting costs. As discussed in Section 3.2.2, introducing unobserved aggregate shocks may help explain the fat tails in the distribution of the margin of victory. It is therefore plausible that extending the model to incorporate such shocks could further improve its predictions and help close the remaining gap for large elections. We leave a complete analysis of this possibility to future research.

5.2.2 What are we learning about the size and number of groups?

The estimates presented in the preceding subsection provide evidence for grouped turnout consistent with the mechanism underlying the Curse of the Silent Majority. While the estimates do not establish that an organized minority defeated the alternative preferred by a majority of the electorate, they assess the quantitative relevance of a group-based turnout mechanism capable of generating such outcomes. Within this framework, we ask two questions: How large is the population in organized groups relative to independents, and what is the nature of these social groups?

The model identifies expected strategic vote shares, $\theta_{s,i} = m_i\lambda_{s,i}/N_i = \alpha_{s,i}\beta_{s,i}$, rather than the organized population share $\alpha_{s,i}$ and its participation probability $\beta_{s,i}$ separately. Because $\beta_{s,i} \in [0, 1]$, the fitted share $\widehat{m}_i\widehat{\lambda}_{s,i}/N_i$ is a lower bound on the share of the electorate organized on side s. Figure 3 plots these bounds using separate fits for each electorate-size stratum.

Figure 3 nevertheless shows that the data do not require a large organized share on either side of the election. In medium-sized jurisdictions, for instance, about 16% of the observations for side R and 26% for side L can be rationalized with no more than 10% of the population organized on that side, and essentially all observations can be rationalized with no more than 25% organized on either side. The result is even starker in large jurisdictions. For each side, more than 30% of the observations can be rationalized with no more than 10% of the population organized on that side, and essentially all observations with no more than 22%. These findings are striking in light of Propositions 3 and 4, since they show that relatively small organized subsets on either side of the electorate may be sufficient to drive the election. Of course, setting the organized share on each side at its lower bound is an extreme scenario, just as assuming that the entire population is organized is an opposite extreme. The most likely scenario is probably somewhere in between. Still, the estimates leave ample room for organized minorities to determine the electoral outcome, a factor that should be considered when evaluating the welfare implications of elections.

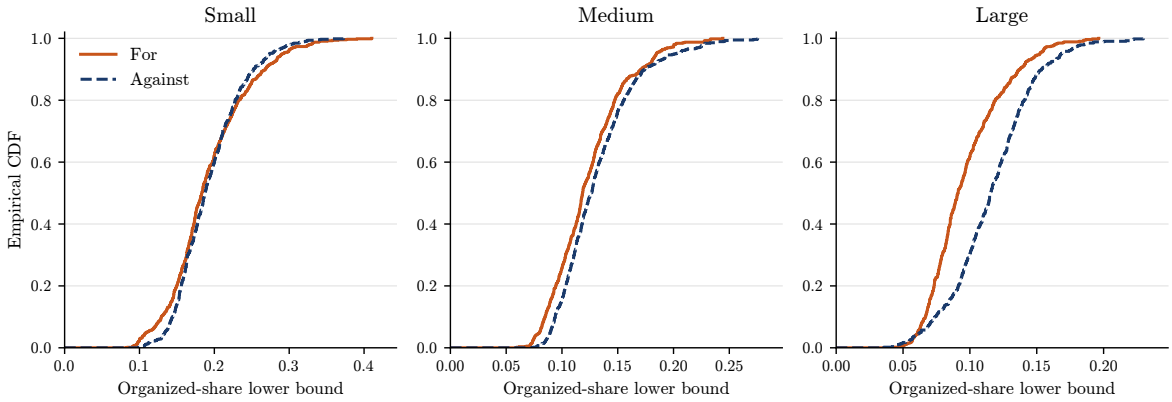


Figure 3: Distribution of the lower bound on the organized share of the electorate by electorate-size stratum (small, medium, and large jurisdictions). Each panel reports the empirical distribution of the election-level bound for R and L.

Consider now the question of what these groups are. First, how large are the groups, and how does their size depend on the nature of the district? Our theory suggests that groups should be sufficiently small, since they must allow their members to trust one another, share information, and coordinate. Still, in large cities there may be structured organizations, such as unions or nonprofits, that allow for larger groups. But how large?

As said above, the mean estimate of the size of organized groups m in the full sample is 42.3, with a median of 28. The estimated size, however, varies with the size of the electoral district. In small jurisdictions, the separately estimated mean group size is 11.5; it is 31.0 in medium jurisdictions; and it is on average 83.7 in large jurisdictions. It is interesting to note that m is increasing in the size of the jurisdiction, but does not explode in large districts with hundreds of thousands of eligible voters. Thus, the model does not rely on a small number of very large groups to fit the evidence. Figure 4 illustrates the distribution of the estimated group size relative to the electorate, plotting the empirical CDF of $\hat{m}/N$ for the full sample and for separate electorate-size fits. The relative group sizes differ across strata, with smaller group shares in larger jurisdictions. Groups are small relative to the electorate: the estimated share exceeds 0.05 of eligible voters in 0.17% of full-sample elections and never exceeds 0.20, and its full-sample median is 0.012. Moreover, although m rises with the size of the jurisdiction, the estimated *share* falls, averaging 0.021 in small jurisdictions, 0.011 in medium ones, and 0.006 in large ones. In absolute terms, m rarely becomes large: 95% of districts are associated with an m no larger than 119. Even in the largest jurisdiction in the sample, Garland, with more than 156,000 eligible voters, the estimated group size is around 246—less than 0.3 percent of the electorate. The fact that the model predicts group sizes that remain small relative to the electorate is consistent with our interpretation of groups as social gatherings

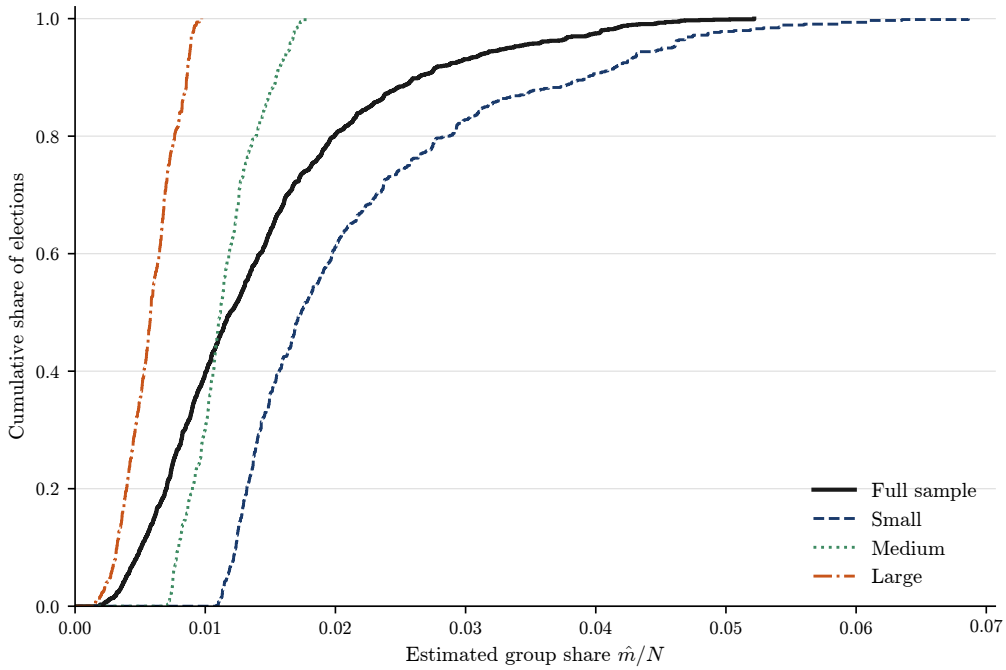


Figure 4: Empirical CDF of the estimated group share $\hat{m}/N$, using the full-sample fit and separate fits within each electorate-size stratum.

in which citizens can communicate, coordinate, and hold one another accountable. The fact that larger jurisdictions are associated with larger groups is consistent with the idea that they are likely to contain more organizations capable of coordinating citizens, such as unionized factories, larger churches, and nonprofit associations.

The estimated group share declines with electorate size, and the city coefficient is negative and statistically significant (Table 7). Thus fitted groups grow less than proportionally with the electorate and are smaller in cities than in county or justice-precinct elections, conditional on the included covariates. These patterns describe differences in effective coordination units; they do not by themselves identify the strength of social ties.

5.2.3 Comparison to other voting models

The preceding comparison restricts the grouped model to $m = 1$. We now compare it with three non-nested alternatives using their probabilities of the same observed vote totals. The models have different mechanisms and observation laws, so the comparison asks which assigns higher likelihood to the common data. The alternatives are:[39]

[39]We do not include the "follow the leader" model of Shachar and Nalebuff (1999), because its driving variable has no counterpart in referenda: turnout responds to unobserved leader effort, and effort is identified from the value of winning each state in a cross-district contest—the state's chance of being pivotal in the Electoral College. Appendix III.1 documents the alternative-model implementations, explains why this effort model cannot be transported

- **Pivotal voter model in Coate et al. (2008)**. We have already discussed this model in Section 5.1. It is the closest benchmark to our restricted model with $m = 1$. We estimate it as in the original CCM paper, updating only the data. This produces a likelihood function defined on the same space of events as our model, namely the space of realized votes in each district.

- **Expressive voting model in Coate et al. (2008)**. As in the Pivotal Voter model, the number of voters supporting liberalization in this voting model is a binomial random variable, and voters have a private cost of voting, c, drawn from a uniform distribution. The substantive difference from the Pivotal Voter model is that this model postulates that a citizen who favors option j votes for j if the voter's realized cost of voting is less than or equal to $\alpha^* v_j$, that is, a fraction α^* of the value v_j of having option j win. In this context, pivotality and other strategic considerations play no role. Here, too, we estimate the model as in the original CCM paper, updating only the data. Again, the likelihood function in this model is defined on the same space of events as our model, namely realized votes.

- **Rule-utilitarian model in Coate and Conlin (2004)**. In this model, each citizen selects the strategy that maximizes the expected welfare of her side. Citizens know the distribution of the share of supporters for liberalizations μ (a Beta distribution whose parameters depend on district characteristics) but not its realization, so each side's cost cutoff is chosen *ex ante*, before the relative group sizes are known; conditional on the realized μ, the model has a continuum of citizens, and the fractions voting for and against are deterministic functions of μ. We re-estimate the model on the full dataset by maximum likelihood with the same functional forms and the same share-density likelihood as the original paper. To place the fitted model on the same count scale as the alternatives, we compute the count likelihood and the simulated moments by integrating over the fitted distribution of the support share using quadrature: this yields the model's finite-count predictive distribution conditional on fixing the common benefit/cost shock at its median. The common shock to benefits and costs that enters their likelihood is held at its median, so that no model in the comparison receives district-level unobserved heterogeneity beyond what its mechanism generates.

Table 8 reports, for each model, the count log likelihood, the implied average turnout, and the simulated dispersion and tail probabilities of turnout and the margin of victory. Among the specifications considered, the likelihood ranking is unambiguous. The grouped model attains a

directly to single-district elections, and reports estimates of a referendum adaptation in which equilibrium effort is proxied by its determinants (Table 16).

count log likelihood of −29,170, compared with −487,407 for the $m = 1$ benchmark, −1,450,777 for CCM, −625,112 for the group rule-utilitarian model, and −718,679 for expressive voting.

Why does the model with groups perform better than these alternative models? Consider first CCM's pivotal-voter and the expressive models. These approaches make individual voting decisions conditionally independent across large numbers of voters. As a result, the law of large numbers forces the predicted distribution of the margin of victory to collapse. For instance, CCM predicts $\Pr(|M| > 0.10) = 0.036$, the $m = 1$ benchmark predicts 0.006, and the expressive model predicts 0.008, all against an empirical frequency of 0.247. Figure 5 illustrates this problem for the CCM pivotal-voter model (a similar figure for the expressive model is presented in the appendix). Each panel corresponds to a single election chosen to approximately span the size quantiles. The empirical outcome in the district is shown as a point in the two-dimensional space of turnout and margin of victory, marked by a star, while simulated outcomes from the grouped model and from CCM are shown in different colors. The CCM model produces a very narrow set of predictions, concentrated around a small region that typically misses the true realization. The grouped model generates a wider range of outcomes, but it is typically better centered on the true outcome.

The story is different for the group rule-utilitarian model. Table 8 shows that this model can explain the margin of victory about as well as, and in some cases better than, the grouped model. For instance, it predicts $\Pr(|M| > 0.05) = 0.513$, versus an empirical value of 0.485; the corresponding value in the grouped model is 0.401. It also predicts $\Pr(|M| > 0.10) = 0.220$, versus an empirical value of 0.247; the corresponding value in the grouped model is 0.141.

Still, the group rule-utilitarian model performs much worse than the grouped model in terms of the likelihood function: −625,112 versus −29,170. The reason is clear from the way the model generates the distribution of outcomes. With a continuum of voters and symmetric strategies, the group rule-utilitarian model would generate a degenerate distribution of turnout unless one introduced an aggregate shock to the support share μ.[40] The model introduces this aggregate uncertainty by making the share of support in district i, μ_i, stochastic and Beta-distributed. Conditional on μ_i, turnout rates in the continuum model are deterministic. Hence, as μ_i varies, the predicted outcome moves along the one-dimensional locus

$$\big(T_i(\mu_i), M_i(\mu_i)\big) = \big((1-\mu_i)\sigma_{R,i} + \mu_i\sigma_{L,i},\ (1-\mu_i)\sigma_{R,i} - \mu_i\sigma_{L,i}\big). \tag{19}$$

The locus of possible outcomes generated by the model with a continuum of voters is the straight line segment with slope $\frac{\sigma_{R,i}+\sigma_{L,i}}{\sigma_{R,i}-\sigma_{L,i}}$ in the (T, M) space, defined by equation (19) as μ_i varies (correspond-

[40] With a continuum of voters of mass μ supporting liberalization who vote with probability σ_s, and a continuum of voters of mass $1-\mu$ opposing liberalization who vote with probability σ_o, turnout is $\mu\sigma_s + (1-\mu)\sigma_o$ and the margin of victory is $(1-\mu)\sigma_o - \mu\sigma_s$ with probability one.

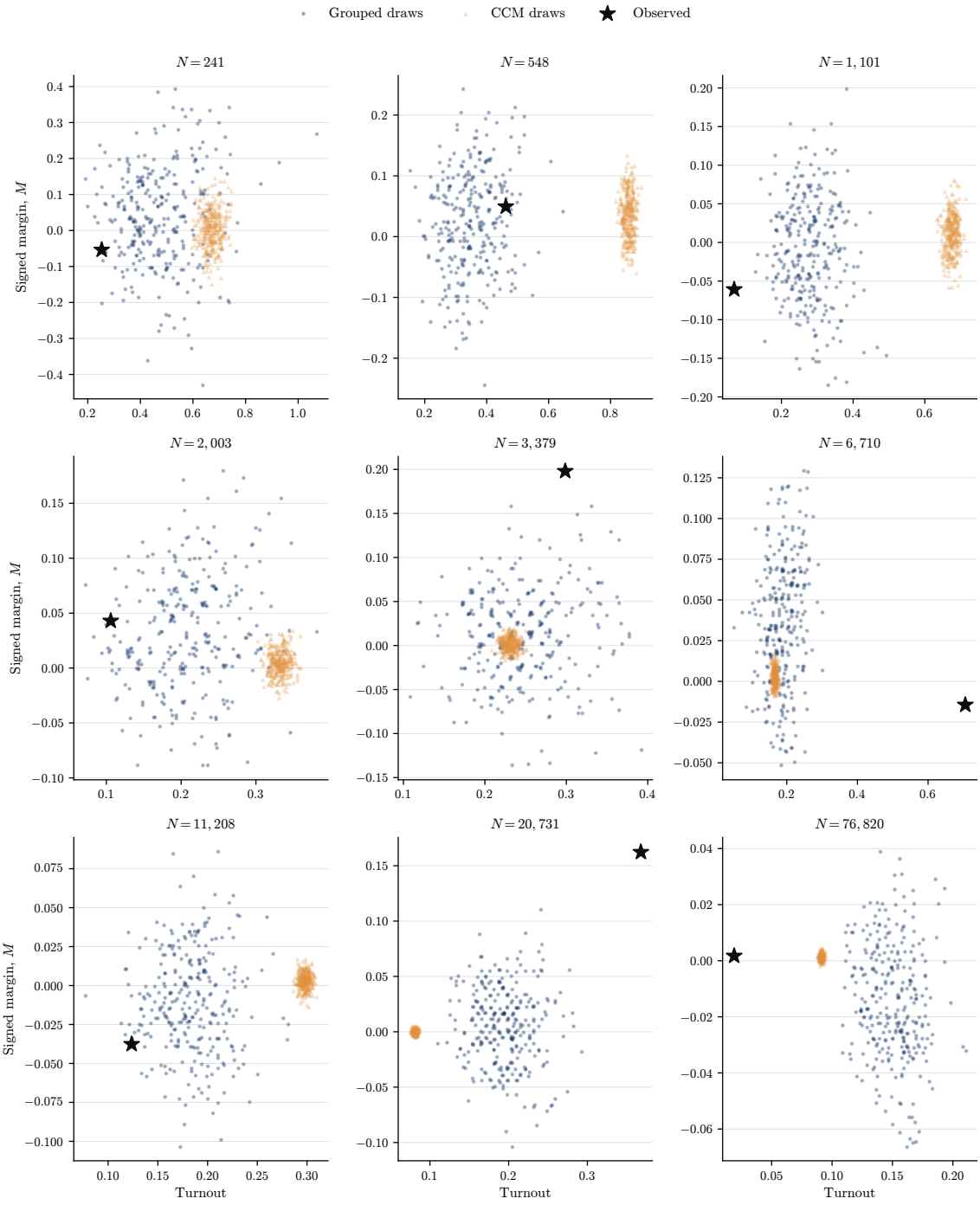


Figure 5: Predictive distributions election by election under the CCM pivotal-voter model, for nine elections at fixed quantiles of the electorate size. Orange triangles: 300 simulated draws from the CCM model; blue circles: 300 draws from the grouped model; star: the observed outcome.

Table 8: Alternative Voting-Behavior Models

Statistic	Data	GRU	Expressive	CCM	$m = 1$	Grouped
Count log likelihood	–	−625,112	−718,679	−1,450,777	−487,407	−29,170
Turnout	0.292	0.241	0.228	0.402	0.299	0.290
SD turnout	0.181	0.097	0.075	0.280	0.104	0.128
SD margin	0.102	0.087	0.040	0.035	0.022	0.073
$\Pr(\lvert M\rvert > .05)$	0.485	0.513	0.216	0.105	0.049	0.401
$\Pr(\lvert M\rvert > .10)$	0.247	0.220	0.008	0.036	0.006	0.141
$\Pr(\lvert M\rvert > .15)$	0.128	0.079	0.000	0.012	0.001	0.052
$\Pr(\lvert M\rvert > .20)$	0.068	0.029	0.000	0.004	0.000	0.020
LR Statistic	–	–	–	–	916,474.26	–
Vuong z	–	10.804	15.502	10.037	15.150	–
Vuong p	–	< 0.001	< 0.001	< 0.001	< 0.001	–
Clarke B	–	1,475	1,618	1,702	1,636	–
Clarke p	–	< 0.001	< 0.001	< 0.001	< 0.001	–

Notes: Full sample of 1,724 elections. GRU follows Coate and Conlin (2004); expressive and CCM follow Coate et al. (2008). Model entries are simulated moments; count log likelihoods evaluate the same observed vote totals. GRU fixes the benefit/cost shock at zero and integrates its fitted Beta support distribution. $M = \tilde{\theta}_R - \tilde{\theta}_L$. The grouped and $m = 1$ models estimate separate indices for $\log \gamma_R$ and $\log \gamma_L$. The LR compares these nested models under the same fixed Skellam noise rule. Vuong statistics use an independent-election normal reference and include the native-estimation influence correction for GRU.

ing to different realizations). In the discrete-voter version we simulate, voters vote independently, so the locus is not exactly a straight line. However, the law of large numbers makes the simulated outcomes very close to such a line. This can be seen in Figure 6, which plots the observed outcome and repeated simulations of both the grouped model and the group rule-utilitarian model, as in Figure 5. The group rule-utilitarian model tends to concentrate predictions along a one-dimensional line that often misses the observed outcome. The grouped model instead generates a less constrained cloud of predictions and is better able to place probability mass around the observed realization.

Table 8 compares election-level log scores on the same observations. For the three non-nested alternatives, the Vuong statistics favor the grouped model and range from 10.037 to 15.502. Their independent-election normal reference gives $p < 0.001$ in each comparison.[41] The Clarke statistic counts the elections in which the grouped model has a higher log score: 1,475 against the group rule-utilitarian model, 1,618 against expressive voting, and 1,702 against CCM, or approximately 86%–99% of the sample. The reported binomial references are nominal election-level sign comparisons. The Vuong and Clarke entries for the nested $m = 1$ model are descriptive diagnostics; the evidence on that restriction uses the likelihood-ratio comparison and fitted-null reference above.

[41] A separate county-block score bootstrap gives $p = 0.00019$ against the group rule-utilitarian model, $p = 0.00008$ against expressive voting, and $p = 0.01858$ against CCM. Thus all three comparisons reject equal expected log scores at the 5% level.

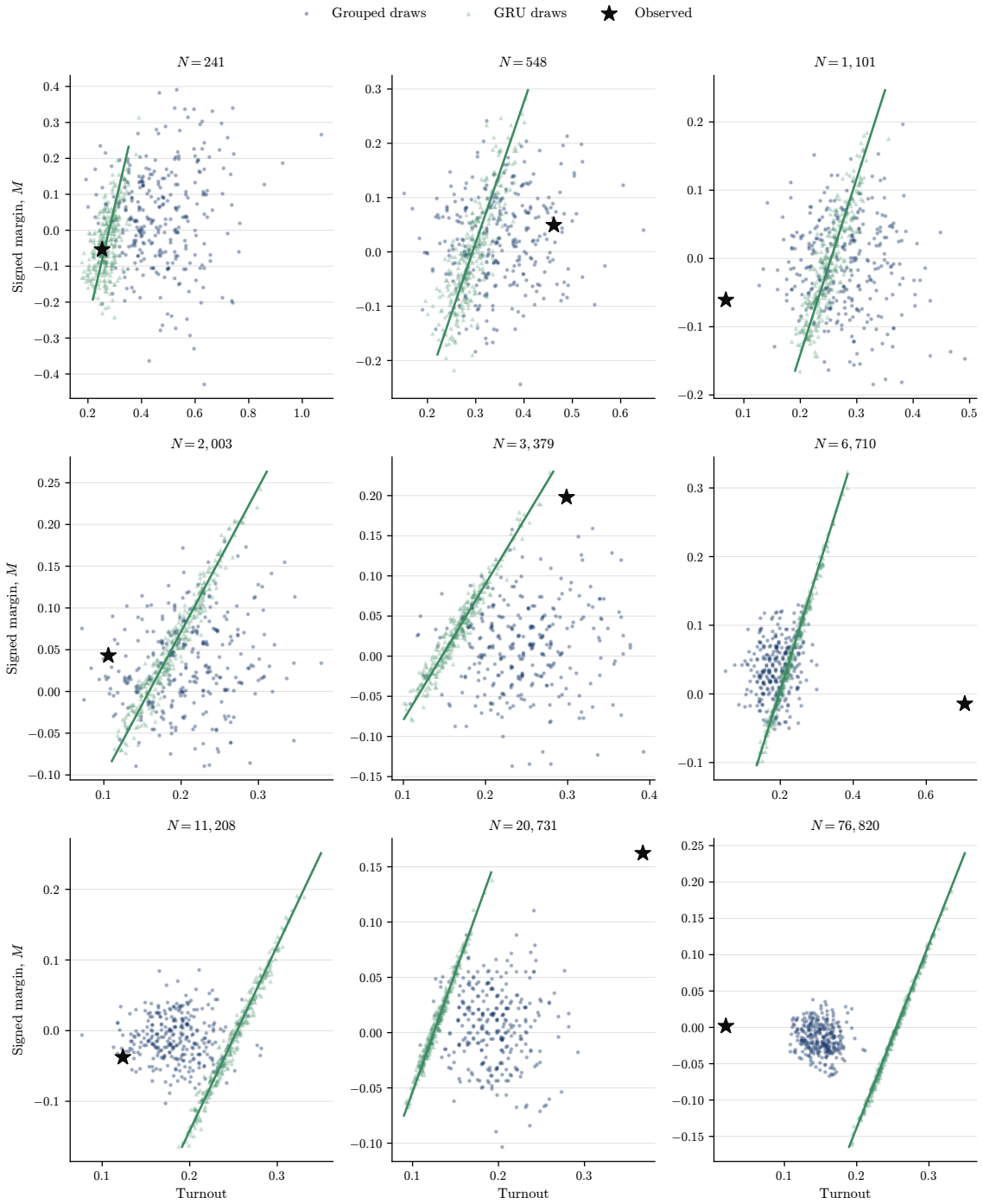


Figure 6: Predictive distributions election by election, for nine elections at fixed quantiles of the electorate size. Green line and triangles: the group rule-utilitarian model's predictive locus and simulated draws; blue circles: the grouped model's simulated draws; star: the observed outcome. As the electorate grows, the rule-utilitarian model's predictive distribution collapses onto its one-dimensional locus, while the grouped model's remains two-dimensional.

5.3 Robustness of the empirical results

Appendix II examines the evidence for grouped voting across periods, electorate-size strata, covariate sets, and sample restrictions. These estimates retain the main mean-zero noise rule, with each Poisson component having mean $0.005N_i$. The grouped model fits better than the corresponding $m = 1$ benchmark in every reported sample comparison (Tables 11 and 14). This includes restricting to first referenda, alternative treatments of justice precincts, and excluding standard November and regular primary dates.

Table 13 examines the sensitivity of the estimates to alternative means of the Poisson noise component in (18), including a constant mean of ten votes that does not vary with electorate size. The grouped model retains its likelihood advantage under every reported noise rule. The magnitude of group size is sensitive to this choice, even though mean turnout changes little across the proportional-noise specifications. This is discussed in detail in Appendix II.

5.4 Out of sample validation

One criticism could be that our model over-fits the data in sample while performing poorly out of sample. We validate our predictions out of sample using ten-fold cross-validation by county. Each of the 1,724 elections is held out once, and both models are fitted using only the remaining counties. The grouped model has higher held-out turnout correlation and lower mean absolute error overall, and improves both measures in seven of the ten folds (Appendix Table 15).

The difference is particularly clear for electoral margins. Margins larger than 10% occur in 24.7% of held-out elections, compared with 13.9% under the grouped model and 0.6% under $m = 1$. Thus groups substantially improve out-of-sample dispersion, although the model still understates the observed tails.

6 Discussion of the theoretical model

We discuss in this section two simplifying assumptions made in the basic model analyzed above. We argue that while these assumptions make the analysis more transparent, relaxing them does not eliminate the mechanisms underlying our main results.

6.1 Asymmetries between R and L Groups and Group-size heterogeneity

The basic model assumes, for simplicity, that all groups are the same size on both the R and L sides. Consider instead an extension in which group size differs across sides: groups supporting R have size m_R, whereas groups supporting L have size m_L. More generally and realistically, m_R and m_L may be random variables with different supports and distributions, so that each side contains groups

of heterogeneous sizes. We argue that, although these natural extensions substantially complicate the analysis, the argument below suggests that they preserve the key pivot-probability mechanism underlying our results.

Consider first the case of deterministic but asymmetric group sizes, in which L- and R-groups have respective sizes $m_L > 1$, $m_R > 1$ with $m_R > m_L$ (the case with $m_R < m_L$ is analogous). Specifically, consider a group-only equilibrium in which both types of groups mix with interior probabilities. Would it still be rational for independent voters to abstain in large elections even if they have higher values and they are more numerous? In such an equilibrium, an L-group is indifferent between voting and abstaining, so $v_L Piv_L = c$, where Piv_L is an L-group's pivot probability. An L-group is pivotal in two events: when there is a tie; and when L is at most m_L votes below R. To specify the probabilities of these events, define:

$$Piv^L_{-l}(m_L, m_R) = \Pr\{i, j \,|\, im_L + l = jm_R\},$$

where i and j here are respectively the realized number of L- and R-groups, so this is just the probability that an L-group is l votes below. The pivot probability associated to a tie is $\frac{1}{2}Piv^L_0(m_L, m_R)$, since when an L-group breaks a tie, L's winning probability increases by $1/2$. The pivot probability associated to the other event is:

$$\sum_{l=1}^{m_L-1} Piv^L_{-l}(m_L, m_R) + \frac{1}{2}Piv^L_{-m_L}(m_L, m_R),$$

since when L is behind by between 1 and $m_L - 1$ votes, the group turns a loss into a victory, whereas when L is exactly m_L votes behind, it creates a tie. By the L-group's indifference condition, we therefore have:

$$c = v_L \left(\begin{array}{c} \frac{1}{2}Piv^L_0(m_L, m_R) \\ + \sum_{l=1}^{m_L-1} Piv^L_{-l}(m_L, m_R) + \frac{1}{2}Piv^L_{-m_L}(m_L, m_R) \end{array} \right) = v_L Piv_L. \tag{20}$$

The right-hand side of this equilibrium condition is the equivalent of the right-hand side of (2). The pivotal event associated to the right-hand side of (20) is however more complicated than the expression when $m_L = m_R = m$ since $\sum_{l=1}^{m_L-1} Piv^L_{-l}(m_L, m_R)$ is always zero when only groups of size m vote in block. When group sizes are heterogeneous, even if only groups vote in block, $Piv^L_{-l}(m_L, m_R)$ may be strictly positive for some $l \in \{1, ..., m_L - 1\}$. For example, when $m_L = 3$ and $m_R = 4$, then $Piv^L_{-1}(m_L, m_R)$ is equal to the probability that, for some nonnegative integer k, we have $i = 4k + 1$ groups supporting L, $j = 3k + 1$ groups supporting R.

The important point here is that the pivot probability for an independent voter is always strictly smaller than that for an L-group. To see this note that the pivot probability for an independent is $Piv^I = \frac{1}{2}Piv^L_0(m_L, m_R) + \frac{1}{2}Piv^L_{-1}(m_L, m_R)$: the first term is the pivot probability when the groups

are tied; and the second term is the pivot probability when L is one vote below R. As for L-groups, when $m_R = m_L = m$, it is never the case that $Piv^L_{-1}(m_L, m_R) > 0$ when only groups vote, but, as mentioned above, $Piv^L_{-1}(m_L, m_R)$ may be strictly positive when $m_R > m_L$. Still, we have:

$$Piv^I = Piv^L - \left[\begin{array}{c} \frac{1}{2} Piv^L_{-1}(m_L, m_R) \\ + \sum_{l=2}^{m_L - 1} Piv^L_{-l}(m_L, m_R) + \frac{1}{2} Piv^L_{-m_L}(m_L, m_R) \end{array} \right] < Piv^L$$

where we stress that the last inequality is strict. To see this it is sufficient to note that even if both $\frac{1}{2} Piv^L_{-1}(m_L, m_R)$ and $\sum_{l=2}^{m_L - 1} Piv^L_{-l}(m_L, m_R)$ are zero, we always must have $Piv^L_{-m_L}(m_L, m_R) > 0$. To show this strict inequality, we need to show that there is always a number i^* of L-groups and a number j^* of R-groups with positive probability such that $i^* \cdot m_L + m_L = j^* \cdot m_R$. A positive probability pair with this property is $i^* = m_R - 1$ and $j^* = m_L$, which gives: $i^* \cdot m_L + m_L = j^* \cdot m_R = m_R m_L$.

This comparison shows that asymmetric block sizes are compatible with abstention by independents even when they are more motivated than L-group members, provided that v_I / v_L is not too large. Specifically, if the L-group is indifferent between voting and abstaining, abstention by independents is optimal whenever $v_I / v_L \leq Piv^L / Piv^I$. A complete characterization of this bound requires solving for the equilibrium participation rates under asymmetric group sizes.

The case with stochastic group sizes is undoubtedly more complex but is conceptually similar. In such an extension, turnout on each side results from a combination of groups. This makes pivotal events more cumbersome, both notationally and computationally, as the number of possible group-type combinations that can produce a given turnout level grows exponentially with the number of group types. Nevertheless, this intuitive argument suggests that even in this more complex setting, whenever an individual voter is pivotal, an L-group with size $m_L > 1$ is also pivotal and generally has a strictly higher probability of altering the outcome in its favor, because it is likely to be pivotal in a larger set of events in which an individual is not pivotal. This pivot-probability advantage preserves the mechanism identified above: for any marginal group type, abstention by independents remains optimal provided their higher valuation is not large enough to offset the group's greater pivot probability. A full characterization requires solving for the participation rates of the different group-size types.

6.2 Individual Rationality and incentives for block voting

So far, groups are exogenous and obedience is assumed, so they act as unitary entities. Membership may have intrinsic value, as in religious groups or sports clubs, while groups may enforce compliance through social sanctions such as ostracism. This raises two questions: how valuable must membership be to make participation individually rational, and how much pressure must groups

be able to exert to ensure obedience? In this section, we show that the equilibria constructed in Propositions 2 and 3 are individually rational even if no intrinsic value for participation is assumed. We also show that a punishment strictly below the voting cost c is sufficient to guarantee obedience by group members.

Individual rationality for a group means that its members weakly prefer to remain in the group they were assigned to by nature at the start of the game, rather than to exit and join the ranks of independents. To see that this property is satisfied in equilibrium, consider what happens if a member i decides to quit a group supporting candidate j. Holding the strategies of players outside the original group fixed, after i exits, both i and the remaining group of $m-1$ members strictly prefer to abstain, regardless of the other's voting decision. To see this, let p_0 be the probability that the other groups tie and $p_{-1}^{(j)}$ the probability that they leave candidate j exactly one block behind. Define $a_j = v_j p_0/2$ and $b_j = v_j p_{-1}^{(j)}/2$. The original group's indifference condition gives $a_j + b_j = c$, with both terms strictly positive. After exit, either the individual or the residual group, acting alone, can only break a tie; its gross benefit from voting is therefore a_j, measured per member for the residual group. When the other votes, its additional votes can only create a tie from a one-block deficit, giving gross benefit b_j. Against any mixed action by the other, the benefit is a convex combination of a_j and b_j, and hence strictly below c. Thus both strictly prefer abstention.

Given this, if i quits the group, the expected numbers of other L- and R-groups remain $\alpha_L N/m$ and $\alpha_R N/m$, respectively; the only change is that i's former group now has $m-1$ members and optimally abstains. Player i is pivotal only when the groups tie, thus s/he will find it strictly optimal to abstain, just as all the other independents. Since the original group is indifferent between voting and abstaining, its members' equilibrium utility equals their utility when the group abstains. This implies that a player who quits a group receives the same expected utility as a member of a group: the utility of a nonvoter who expects $\alpha_L N/m$ groups for L each voting in block with probability β_L, and $\alpha_R N/m$ groups for R each voting in block with probability β_R (where β_L and β_R are the equilibrium strategies). Remaining in a group is therefore individually rational even if the group provides no intrinsic benefit from participation. The same argument applies to the interior equilibrium constructed in Proposition 3.

Consider now Obedience. This property means that a member of a group finds it optimal to obey the group's recommendation: to vote together with all other members when the group's simple strategy calls for voting, and to abstain otherwise.

In the equilibrium of Proposition 2, and more generally in any equilibrium in which independents abstain and both types of groups mix with interior probability, we have:

$$c = v_j \mathcal{P}^j(\beta) \simeq \frac{v_j e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2 \frac{N}{m}}}{4\left(\pi \frac{N}{m}\sqrt{\alpha_R\alpha_L\beta_R\beta_L}\right)^{\frac{1}{2}}} \left[1 + \left(\frac{\alpha_{-j}\beta_{-j}}{\alpha_j\beta_j}\right)^{1/2}\right] \tag{21}$$

where the right-hand side is the group's expected instrumental benefit from voting when it behaves as a block. For the displayed pivotality formulas, we use Myerson's approximation. A member of group j who votes when asked to abstain is pivotal only when the other groups tie, which is the first term in (21). By the exact decomposition above, this benefit is $a_j < c$, so the instrumental benefit of voting when asked to abstain is strictly lower than the cost of voting. Any member who is ordered to abstain, therefore, finds it strictly optimal to obey even if there is no punishment $\pi^j \geq 0$ in case of disobedience.

Consider now the incentives to vote when asked by the group. The expected benefit of a group j is given by the right-hand side of (21) because a group of type j is pivotal when the other groups are tied or when j is one group below $-j$. A member i of group j, however, is pivotal only when j is one group below: if all members other than i obey the order to vote and i deviates, the group still wins with probability 1 when the other groups are tied. Assume that not voting when asked implies a penalty π^j in group j. Given obedience by the other members, i is willing to vote when asked if and only if $b_j + \pi^j \geq c$. Under Myerson's approximation, this condition becomes:

$$\frac{v_j e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2 \frac{N}{m}}}{4\left(\pi \frac{N}{m}\sqrt{\alpha_R\alpha_L\beta_R\beta_L}\right)^{\frac{1}{2}}} \left(\frac{\alpha_{-j}\beta_{-j}}{\alpha_j\beta_j}\right)^{1/2} + \pi^j \geq c \tag{22}$$

Using the notation introduced above, a member who is asked to abstain has a gross voting benefit $a_j < c$, while a member asked to vote has a gross voting benefit b_j, given obedience by the other members. The exact minimum penalty for not voting when asked is therefore

$$\underline{\pi}^j = c - b_j = a_j = \frac{v_j p_0}{2} \in (0, c).$$

At this minimum penalty, voting when asked is weakly optimal. Any slightly larger penalty, still below c, makes it strictly optimal.

Using (5), we can approximately rewrite the first term on the left-hand side of (22) as $cv_{-j}/(v_R + v_L)$, and obtain the following lower bound for π^j:

$$\pi^j \geq \underline{\pi}^j \simeq \frac{v_j}{v_R + v_L} c.$$

The exact threshold shows that the required punishment for a group to guarantee obedience is positive but always strictly smaller than the cost of voting, a quantity that is commonly assumed to be small. An immediate implication of the argument presented above, moreover, is that, under the same approximation, if enforcement is given, then for any π^l, with $l = R, L$, the equilibrium is obedient whenever $c < [(v_R + v_L)/v_R] \cdot \min \pi^j$.

7 Conclusion

In this paper, we have presented a model of elections in which part of the electorate consists of independent voters, as in standard models, while the remainder is organized into groups. These groups allow their members to communicate and coordinate their actions, enabling them to employ correlated strategies. Examples include local branches of political parties and labor unions, as well as social organizations that are not explicitly political, such as churches, sports clubs, extended families, and social-media networks.

We have established two closely related sets of theoretical results. First, even when organized groups comprise only a small fraction of the electorate, they may exert a disproportionate influence on electoral outcomes. Under weak conditions, sufficiently large elections admit an equilibrium in which independent voters find it optimal to abstain, leaving the electoral outcome entirely in the hands of the organized minority. Remarkably, this occurs even when independent voters are more numerous than group members and more motivated, in the sense that they have stronger instrumental preferences over the outcome. Separately, when independents' valuation is sufficiently close to that of the L-groups, they abstain in every equilibrium of sufficiently large elections. We call this phenomenon the *Curse of the Silent Majority*.

Second, when turnout is generated by groups whose members vote as blocks, the model helps reconcile voting theory with empirical evidence. In particular, it can account for a phenomenon that existing models have considerable difficulty explaining: the frequent occurrence of large margins of victory.

To demonstrate the empirical relevance of these theoretical findings, we have also structurally estimated the model. Because the standard pivotal-voter model is nested within our framework, the estimation permits a direct comparison between the two, as well as with several other (non-nested) models proposed in the literature. The results provide strong evidence that incorporating groups substantially improves the model's ability to reconcile voting theory with the data and produces a better fit than the alternative approaches we consider. These empirical gains represent a quantitatively meaningful yet partial explanation of large electoral margins. The estimation also yields new insights into the social organizations that influence electoral outcomes.

We view these contributions as a first step toward a broader study of how social groups shape elections. To obtain closed-form and readily interpretable results, we have modeled groups in a deliberately simple way. In particular, we assume that voters are either independent or belong to groups of a common size m, thereby ruling out heterogeneity across groups (both in terms of their sizes and motivations of their members). We do not model communication within groups directly, instead limiting attention to simple block strategies. Finally, we assume perfect coordination within

each group but complete independence across groups. The social structure underlying real-world elections is clearly much richer. Voters may belong to a variety of groups that differ in size, composition, and the intensity of their members' preferences. A more general framework could explicitly model communication and potential conflicts of interest within groups, as well as different degrees of coordination both within and across groups. For example, local churches may retain considerable autonomy while still being partially coordinated through a broader denominational hierarchy. Incorporating these features would produce a more nuanced and complex environment, but it would also open several promising avenues for future research. We plan to pursue some of these extensions in subsequent work.

I Appendix A: Proofs

I.1 Proof of Lemma 1

Lemma 1: In a large election without organized group coordination, $\sigma \in (0,1)$ and $\beta_L = 0$ in any equilibrium. That is, independents vote for L with a strictly positive probability less than 1, and L-type group members always abstain.

Proof: We proceed in two steps. First we prove that under the assumptions of the lemma, $\sigma \in (0,1)$ in any equilibrium with uncoordinated groups. We then prove that $\beta_L = 0$.

Step 1. First, it is easy to see that for a given $c > 0$, we cannot have $\sigma = 1$ for large enough N. If it were the case that $\sigma = 1$, then L would win with probability approaching one as $N \to \infty$, for any possible turnout probabilities of β_R since $\alpha_I > \alpha_R$ and the voters in groups of type L never vote for R. Hence an independent voter's best reply for large N is to abstain, a contradiction.

We now prove that we cannot have an equilibrium with $\sigma = 0$. If this were the case, we would have:

$$\frac{c}{v_L} > \frac{c}{v_I} \geq \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\beta_L\alpha_L N}(\beta_L\alpha_L N)^k}{k!}\right)\left(\frac{e^{-\beta_R\alpha_R N}(\beta_R\alpha_R N)^k}{k!}\right)\left(1+\frac{\beta_R\alpha_R N}{k+1}\right) \tag{23}$$

since $v_L < v_I$. It follows from (23) that $\beta_L = 0$. The key observation here is that the independents and the members of the groups for L are pivotal in the same event, and so have the same pivot probability (i.e. to make or break a tie), which is given by the expression on the right in (23). It is then easy to see that we also cannot have an equilibrium in which $\sigma = 0$, $\beta_L = 0$ and $\beta_R = 0$, otherwise the optimal reply for every voter would be to vote, since $c < v_i/2$ by assumption.

Therefore, to prove that $\sigma \in (0,1)$, we only need to show that we cannot have an equilibrium in which $\sigma = 0$, $\beta_L = 0$ and $\beta_R > 0$. If there were such an equilibrium, we would have:

$$\frac{c}{v_R} \leq \frac{1}{2}e^{-\beta_R\alpha_R N} < \frac{1}{2}e^{-\beta_R\alpha_R N}\left(1+\beta_R\alpha_R N\right) \leq \frac{c}{v_I}$$

a contradiction, since $v_R < v_I$. The first inequality follows from the optimality of R voters to vote given $\sigma = 0$, $\beta_L = 0$, and the last inequality follows from the optimality of independent voters to abstain given $\sigma = 0$, $\beta_L = 0$. Hence, $\sigma \in (0,1)$.

Step 2. If $\sigma \in (0,1)$, we must have:

$$\frac{c}{v_L} > \frac{c}{v_I} = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-(\sigma\alpha_I+\beta_L\alpha_L)N}\left[\left(\sigma\alpha_I+\beta_L\alpha_L\right)N\right]^k}{k!}\right)\left(\frac{e^{-\beta_R\alpha_R N}(\beta_R\alpha_R N)^k}{k!}\right)\left(1+\frac{\beta_R\alpha_R N}{k+1}\right)$$

so $\beta_L = 0$. It follows that we must have $\sigma \in (0,1)$, $\beta_L = 0$ and $\beta_R \in [0,1]$. ■

I.2 Proof of Proposition 1

Proposition 1: In a large election without organized groups, for any $\delta > 0$, there is a $c(\delta)$ such that for $c < c(\delta)$, we have $P^L(c) \geq 1 - \delta$.

Proof: Consider any sequence $(c_t, N_t)_{t=1}^{\infty}$ such that $c_t > 0$, $c_t \to 0$, and $N_t \to \infty$, and any corresponding sequence of equilibrium strategies. If $\sigma_t = 1$ along a subsequence, then independent expected turnout is $\alpha_I N_t$, whereas R-turnout is at most $\alpha_R N_t$. Because $\alpha_I > \alpha_R$, candidate L wins with probability approaching one along that subsequence. We may therefore restrict attention to a subsequence on which $\sigma_t < 1$. We start from a preliminary lemma.

Lemma 1.1 Consider any sequence $(c_t, N_t)_{t=1}^{\infty}$ such that $c_t > 0$, $\lim_{t\to\infty} c_t = 0$, and $\lim_{t\to\infty} N_t = \infty$, and any corresponding sequence of equilibrium strategies $(\sigma_t, \beta_{R,t})_{t=1}^{\infty}$. For any $T^L > 0$, there is a $\bar{t}$ such that $T_t^L \equiv \sigma_t \alpha_I N_t > T^L$ for all $t > \bar{t}$.

Proof. We explicitly begin with the possibility that $\sigma_t = 1$. Along any subsequence on which $\sigma_t = 1$, we have $T_t^L = \alpha_I N_t \to \infty$. Thus, if the lemma were false, there would be a subsequence along which T_t^L is bounded and, after omitting finitely many terms, $\sigma_t < 1$. Since $\sigma_t = 0$ is impossible, we then have $\sigma_t \in (0, 1)$. Moreover, the argument in Lemma 1 implies $\beta_{L,t} = 0$.

If R-types vote with zero probability along a further subsequence, then only independents vote and independent indifference gives $c_t/v_I = e^{-\sigma_t \alpha_I N_t}/2$. Therefore, $T_t^L = \log(v_I/(2c_t)) \to \infty$, contrary to boundedness. Hence, on a putative bounded subsequence, we may pass to a further subsequence on which R-types vote with positive probability.

We next prove that $\sigma_t \alpha_I N_t > \alpha_R \beta_{R,t} N_t$. Since $\sigma_t \in (0, 1)$ and $\beta_{L,t} = 0$, we must have:

$$\frac{c_t}{v_I} = \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{e^{-\sigma_t \alpha_I N_t} (\sigma_t \alpha_I N_t)^k}{k!} \right) \left(\frac{e^{-\alpha_R \beta_{R,t} N_t} (\alpha_R \beta_{R,t} N_t)^k}{k!} \right) \left(1 + \frac{\alpha_R \beta_{R,t} N_t}{k+1} \right) \tag{24}$$

$$\frac{c_t}{v_R} \leq \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{e^{-\sigma_t \alpha_I N_t} (\sigma_t \alpha_I N_t)^k}{k!} \right) \left(\frac{e^{-\alpha_R \beta_{R,t} N_t} (\alpha_R \beta_{R,t} N_t)^k}{k!} \right) \left(1 + \frac{\sigma_t \alpha_I N_t}{k+1} \right). \tag{25}$$

The right-hand side of equation (24) is the pivot probability for an independent: that is, the probability that a vote by an independent changes the outcome of the election in favor of L. Similarly, the right-hand side of inequality (25) is the pivot probability for an R group member: that is, the probability that a vote by an R group member changes the outcome of the election in favor of R. If, to the contrary, $\sigma_t \alpha_I N_t \leq \alpha_R \beta_{R,t} N_t$, then every summand in (24) is at least the corresponding summand in (25). It would follow that $c_t/v_I \geq c_t/v_R$, a contradiction since $v_I > v_R$. Hence $T_t^L = \sigma_t \alpha_I N_t > \alpha_R \beta_{R,t} N_t = T_t^R$.

We can now complete the proof of Lemma 1.1. On the putative bounded subsequence, pass to a further subsequence such that $T_t^L \to T_\infty^L < \infty$. Since $T_t^R < T_t^L$, pass to a further subsequence

such that $T_t^R \to T_\infty^R < \infty$. Thus turnout from both independents and R types is bounded in the limit. From (24), retaining only the $k = 0$ term, we have

$$0 = \lim_{t\to\infty} \frac{c_t}{v_I} \geq \frac{1}{2} \lim_{t\to\infty} \left[(1 + \alpha_R \beta_{R,t} N_t)\, e^{-(\alpha_I \sigma_t + \alpha_R \beta_{R,t}) N_t} \right] = \frac{1}{2} \left(1 + T_\infty^R\right) e^{-\left(T_\infty^L + T_\infty^R\right)} > 0,$$

which is a contradiction. We conclude that $T_t^L \to \infty$ as $t \to \infty$. ■

Using Lemma 1.1, we can now prove that, for any sequence c_t, N_t and associated equilibrium strategies $\sigma_t, \beta_{R,t}$ such that $c_t \to 0$ and $N_t \to \infty$, the probability that L wins converges to one; that is, $P^L(\sigma_t, \beta_{R,t}) \to 1$. Expected turnout for L and R is, respectively, $T_t^L = \sigma_t \alpha_I N_t$ and $T_t^R = \alpha_R \beta_{R,t} N_t$. Consider an arbitrary subsequence. Passing to a further subsequence if necessary, either T_t^R is bounded, in which case we may assume that $T_t^R \to T_\infty^R < \infty$, or $T_t^R \to \infty$.

In the first case, Lemma 1.1 gives $T_t^L \to \infty$. If D_t denotes the number of votes for L minus the number of votes for R, then $E[D_t] = T_t^L - T_t^R$ and $\mathrm{Var}(D_t) = T_t^L + T_t^R$. Consequently, $E[D_t]/\sqrt{\mathrm{Var}(D_t)} = (T_t^L - T_t^R)/\sqrt{T_t^L + T_t^R} \to \infty$, so candidate L wins with probability approaching one.

Consider now the second case, in which $T_t^R \to \infty$. Lemma 1.1 also gives $T_t^L \to \infty$. Whether R-types mix or vote with probability one, their best-response condition and the independent indifference condition, together with the limit formulas in Myerson (2000), give

$$c_t \leq v_R \frac{e^{-\left(\sqrt{\beta_{R,t}\alpha_R N_t} - \sqrt{\sigma_t \alpha_I N_t}\right)^2}}{4\left(\pi \sqrt{\beta_{R,t}\alpha_R N_t \cdot \sigma_t \alpha_I N_t}\right)^{\frac{1}{2}}} \left(\frac{(\beta_{R,t}\alpha_R N_t)^{\frac{1}{2}} + (\sigma_t \alpha_I N_t)^{\frac{1}{2}}}{(\beta_{R,t}\alpha_R N_t)^{\frac{1}{2}}} \right) (1 + o(1)) \tag{26}$$

$$c_t = v_I \frac{e^{-\left(\sqrt{\beta_{R,t}\alpha_R N_t} - \sqrt{\sigma_t \alpha_I N_t}\right)^2}}{4\left(\pi \sqrt{\beta_{R,t}\alpha_R N_t \cdot \sigma_t \alpha_I N_t}\right)^{\frac{1}{2}}} \left(\frac{(\beta_{R,t}\alpha_R N_t)^{\frac{1}{2}} + (\sigma_t \alpha_I N_t)^{\frac{1}{2}}}{(\sigma_t \alpha_I N_t)^{\frac{1}{2}}} \right) (1 + o(1)). \tag{27}$$

These asymptotic formulas are valid with relative error $o(1)$ because turnout on both sides diverges. Equations (26)–(27) imply $v_I/v_R \leq \sqrt{T_t^L/T_t^R}\,(1 + o(1))$. More precisely, let $\mathcal{P}_R^t, \mathcal{P}_I^t$ denote, respectively, the pivotal probabilities of an R-type voter and an independent voter in election t. The exact Skellam formulas imply $\frac{\mathcal{P}_R^t}{\mathcal{P}_I^t} = \frac{1+s_t\rho_t}{1+\rho_t/s_t} = s_t[1 + o(1)]$ uniformly for $s_t \geq 1$, where $s_t = \sqrt{\frac{T_t^L}{T_t^R}} \geq 1$, and $\rho_t = \frac{I_1\left(2\sqrt{T_t^L T_t^R}\right)}{I_0\left(2\sqrt{T_t^L T_t^R}\right)}$ (here, I_k denotes the modified Bessel function of the first kind of order k). Since $T_t^L, T_t^R \to \infty$, the standard large-argument expansion of the modified Bessel functions implies $\rho_t \longrightarrow 1$. We therefore have $\left|\frac{1+s_t\rho_t}{s_t+\rho_t} - 1\right| = \frac{(s_t-1)(1-\rho_t)}{s_t+\rho_t} \leq 1 - \rho_t \longrightarrow 0$. It follows that the relative-$o(1)$ comparison obtained by dividing (26) and (27) is valid uniformly, even if T_t^L/T_t^R is unbounded.

Therefore, $T_t^L/T_t^R \geq (v_I/v_R)^2\,(1 + o(1))$. Since $v_I/v_R > 1$, choose a constant $\eta > 0$ such that $1 + \eta < (v_I/v_R)^2$. For all sufficiently large t, $T_t^L \geq (1 + \eta) T_t^R$. It follows that $T_t^L - T_t^R \geq \eta T_t^R$

and $T_t^L - T_t^R \geq [\eta/(2+\eta)](T_t^L + T_t^R)$. Since $T_t^R \to \infty$, we have $(T_t^L - T_t^R)/\sqrt{T_t^L + T_t^R} \to \infty$. Chebyshev's inequality then implies that candidate L wins with probability approaching one.

We have shown that every subsequence has a further subsequence along which L's winning probability converges to one. Hence, for every sequence c_t, N_t and associated equilibrium strategies such that $c_t \to 0$ and $N_t \to \infty$, we have $P^L(\sigma_t, \beta_{R,t}) \to 1$. It follows that $\lim_{c\to 0} \liminf_{N\to\infty} P_N^L(c) = 1$, so $\lim_{c\to 0} P^L(c) = 1$. To see this, suppose the proposition were false. Since $P^L(c) \leq 1$, there would exist $\varepsilon > 0$ and a sequence $c_t \downarrow 0$, with $c_t < 1/t$, such that $\liminf_{N\to\infty} P_N^L(c_t) \leq 1 - 2\varepsilon$ for every t. By the definition of the lim inf, for each t we could choose $N_t > t$ such that $P_{N_t}^L(c_t) \leq 1 - \varepsilon$. Thus $c_t \to 0$ and $N_t \to \infty$, while $P_{N_t}^L(c_t)$ remains bounded above by $1 - \varepsilon$, contradicting the sequential convergence established above. Therefore, $\lim_{c\downarrow 0} P^L(c) = 1$. It follows that, for every $\delta > 0$, there exists $c(\delta) > 0$ such that $P^L(c) > 1 - \delta$ whenever $0 < c < c(\delta)$. ■

I.3 Proof of Proposition 2

Proposition 2: Assume $c < \bar{c}$. There exists $\Delta^* > 0$ such that if $v_I \leq v_R[1 + \Delta^*]$ then in a large election with organized groups there is an equilibrium in which $\sigma = 0$, $\beta_R, \beta_L \in (0, 1)$, and $\alpha_R\beta_R > \alpha_L\beta_L$. At this equilibrium, the probability that L wins is strictly less than $1/2$.

Proof: We proceed in three steps. In Step 1, we establish existence of an equilibrium in the auxiliary game in which all independent voters abstain and only the organized R-groups and L-groups are active. In Step 2, we show that, for all sufficiently large N, the auxiliary game has an equilibrium in which both group types mix in the interior. Finally, in Step 3, we show that if $v_I \leq v_R(1 + \Delta^*)$, independent voters optimally abstain, so the equilibrium of the modified game is also an equilibrium of the original game.

Step 1. Consider first the auxiliary game in which all independent voters abstain. For an electorate indexed by N, there are on average $\kappa_R^N = \alpha_R N/m$ potential R-groups and $\kappa_L^N = \alpha_L N/m$ potential L-groups. If a group of type j votes with probability β_j, the number of voting groups of that type is Poisson with mean $\lambda_j = \kappa_j^N \beta_j$. The payoff differences between voting and abstaining are continuous in (β_R, β_L). Hence the usual best-response correspondence on $[0, 1]^2$ is nonempty, convex-valued, and upper hemicontinuous, so Kakutani's theorem gives an equilibrium of the auxiliary game for every N.

Step 2. We next show that, for all sufficiently large N, the auxiliary game has an equilibrium in which both group types mix in the interior. In fact, every auxiliary-game equilibrium is interior for all sufficiently large N. Let N_R and N_L be independent Poisson variables with means λ_R and λ_L, and write $p_k = \Pr(N_R - N_L = k)$. Under fair tie breaking, the pivotal probabilities of an R-group and an L-group are $P_R = (p_0 + p_{-1})/2$ and $P_L = (p_0 + p_1)/2$.

For all sufficiently large N, every auxiliary-game equilibrium is interior. First, both types cannot abstain, because their pivotal probabilities would be $1/2$ and $c < \bar{c} < v_L/2 < v_R/2$. Here $\bar{c} < v_L/2$ follows from $x\exp(1-x) < 1$ for $x = v_R/v_L > 1$. Second, full participation by either type is impossible for large N. Indeed, if $M(\lambda) = \sup_{r\geq 0}\Pr(\text{Poisson}(\lambda) = r)$, then $M(\lambda) \longrightarrow 0$ as $\lambda \longrightarrow \infty$, and conditioning on either Poisson count gives $p_k \leq \min\{M(\lambda_R), M(\lambda_L)\}$. If, for example, $\beta_R = 1$, then $\lambda_R = \alpha_R N/m \longrightarrow \infty$, so $P_R \longrightarrow 0$ uniformly in β_L, contradicting the optimality of full participation. The argument for L is identical.

It remains to exclude zero participation by exactly one type. If $\beta_R = 0$, then $\beta_L \in (0,1)$ and the L-groups are indifferent. With $\lambda_R = 0$, we have $P_L = e^{-\lambda_L}/2$ and $P_R = e^{-\lambda_L}(1+\lambda_L)/2$. Thus $v_R P_R > v_L P_L = c$, so an R-group strictly prefers voting. If $\beta_L = 0$, then $\beta_R \in (0,1)$ and $c/v_R = P_R = e^{-\lambda_R}/2$, while $P_L = e^{-\lambda_R}(1+\lambda_R)/2$. Abstention by an L-group would require $v_L P_L \leq c$, and hence $\lambda_R \leq v_R/v_L - 1$. It would follow that $c = (v_R/2)e^{-\lambda_R} \geq (v_R/2)\exp(1 - v_R/v_L) = \bar{c}$, a contradiction.

Let $\bar{N}_0$ be such that all auxiliary-game equilibria are interior for every $N \geq \bar{N}_0$. Fix any $N_0 \geq \bar{N}_0$ and an interior equilibrium $(\beta_R^{N_0}, \beta_L^{N_0})$. Define its turnout intensities by $\lambda_R^* = \alpha_R \beta_R^{N_0} N_0/m$ and $\lambda_L^* = \alpha_L \beta_L^{N_0} N_0/m$. If N_R^* and N_L^* are independent Poisson variables with these means, write $p_k^* = \Pr(N_R^* - N_L^* = k)$. Define the R-group pivotal probability at these intensities by $P_R^* := (p_0^* + p_{-1}^*)/2$. Interior indifference gives $c/v_R = P_R^*$ and $c/v_L = (p_0^* + p_1^*)/2$. Since $v_R > v_L$, we have $p_{-1}^* < p_1^*$. The Skellam identity $p_1^*/p_{-1}^* = \lambda_R^*/\lambda_L^*$ therefore implies $\lambda_R^* > \lambda_L^*$.

For every $N \geq N_0$, set $\beta_R^N = m\lambda_R^*/(\alpha_R N)$ and $\beta_L^N = m\lambda_L^*/(\alpha_L N)$. Equivalently, $\beta_j^N = (N_0/N)\beta_j^{N_0}$, so both probabilities are in $(0,1)$. Poisson thinning keeps the turnout intensities equal to λ_R^* and λ_L^*; hence the pivotal probabilities remain unchanged and both group types remain indifferent. Moreover, $\alpha_R \beta_R^N = m\lambda_R^*/N > m\lambda_L^*/N = \alpha_L \beta_L^N$.

Step 3. It remains to verify the independent voters' incentive to abstain. Because $m > 1$, a single independent vote is pivotal, when all other independents abstain, only if the organized-group vote is tied. Its pivotal probability is therefore $P_I^* = p_0^*/2$. Since both Poisson means are positive, $p_0^* > 0$ and $p_{-1}^* > 0$. Define $\Delta^* = p_{-1}^*/p_0^* > 0$. This constant is independent of N, and $1 + \Delta^* = (p_0^* + p_{-1}^*)/p_0^* = P_R^*/P_I^*$. Consequently, $v_I \leq v_R(1+\Delta^*)$ implies $v_I P_I^* \leq v_R P_R^* = c$, so abstention is a best response for every independent voter. The response is strict if $v_I < v_R(1+\Delta^*)$. Finally, for every $k \geq 1$, the Skellam identity gives $p_k^*/p_{-k}^* = (\lambda_R^*/\lambda_L^*)^k > 1$. Hence $\Pr(N_R^* > N_L^*) > \Pr(N_L^* > N_R^*)$, so after fair tie breaking candidate R wins with probability strictly greater than $1/2$. Taking $\bar{N} = N_0$ completes the proof. ■

I.4 Proof of Proposition 3

Proposition 3: If $v_I < v_R + v_L$ then, in a large economy, for any δ there is a $c(\delta)$ such that for $c < c(\delta)$ there is an equilibrium in which independents abstain and L wins with probability lower than δ.

Proof: Put $q = v_L/v_R \in (0,1)$ and $\bar{c} = (v_R/2)\exp(1 - v_R/v_L)$. For every $c \in (0, \bar{c})$, the auxiliary-game argument used in the proof of Proposition 2 provides an interior equilibrium in which only organized groups vote. Let $x_R(c) > 0$ and $x_L(c) > 0$ denote the expected numbers of voting R-groups and L-groups in one such equilibrium. The exact indifference conditions are $v_R P_R(x_R, x_L) = v_L P_L(x_R, x_L) = c$. Put $t = 2\sqrt{x_R x_L}$, $z = \sqrt{x_L/x_R}$, and $\rho(t) = I_1(t)/I_0(t)$, where I_ν is the modified Bessel function of the first kind. The exact Skellam formulas give

$$P_R = \frac{1}{2} e^{-(x_R + x_L)} \left[I_0(t) + z I_1(t) \right], \qquad P_L = \frac{1}{2} e^{-(x_R + x_L)} \left[I_0(t) + \frac{1}{z} I_1(t) \right].$$

Because $P_R/P_L = v_L/v_R = q$, compatibility of the two indifference conditions is equivalent to $(1 + z\rho(t))/(1 + \rho(t)/z) = q$, or $\rho(t) z^2 + (1-q)z - q\rho(t) = 0$. This quadratic has the unique positive solution $z = h(\rho) := [\sqrt{(1-q)^2 + 4q\rho^2} - (1-q)]/(2\rho)$. Consequently, z is determined by t.

We claim that $t \longrightarrow \infty$ as $c \downarrow 0$. Suppose otherwise. Then there is a sequence $c_\ell \downarrow 0$ along which t_ℓ is bounded. Passing to a subsequence, let $t_\ell \longrightarrow \bar{t} \geq 0$. If $\bar{t} > 0$, continuity and strict positivity of the Bessel functions imply that

$$c_\ell = \frac{v_R}{2} \exp\left[-\frac{t_\ell}{2} \left(z_\ell + z_\ell^{-1} \right) \right] I_0(t_\ell) \left[1 + z_\ell \rho(t_\ell) \right]$$

converges to a strictly positive number, contradicting $c_\ell \longrightarrow 0$.

If $\bar{t} = 0$, the standard small-argument expansions give $\rho(t) \sim t/2$ and $h(\rho(t)) \sim qt/[2(1-q)]$. Since $x_R = t/(2z)$ and $x_L = tz/2$, it follows that $x_R \longrightarrow (1-q)/q$ and $x_L \longrightarrow 0$. Therefore $c \longrightarrow (v_R/2)\exp[-(1-q)/q] = (v_R/2)\exp(1 - v_R/v_L) = \bar{c}$, which again contradicts $c \downarrow 0$. Hence $t \longrightarrow \infty$. The standard large-argument expansion gives $\rho(t) \longrightarrow 1$. The formula for h then implies $z \longrightarrow q$, $x_L(c)/x_R(c) = z^2 \longrightarrow q^2 = (v_L/v_R)^2$, and $x_R(c), x_L(c) \longrightarrow \infty$.

We next verify the independent voters' incentive. When all other independents abstain, an independent voter is pivotal only when the two numbers of voting groups are equal. Thus $P_I(x_R, x_L) = e^{-(x_R + x_L)} I_0(t)/2$. Using the R-group indifference condition,

$$\frac{v_I P_I}{c} = \frac{v_I}{v_R} \frac{I_0(t)}{I_0(t) + z I_1(t)} = \frac{v_I/v_R}{1 + z\rho(t)} \longrightarrow \frac{v_I}{v_R(1+q)} = \frac{v_I}{v_R + v_L} < 1.$$

Therefore, for all sufficiently small c, every independent voter strictly prefers to abstain.

Finally, let $Z(c) = N_R - N_L$, where N_R and N_L are independent Poisson variables with means $x_R(c)$ and $x_L(c)$. Then $E[Z(c)] = x_R(c) - x_L(c)$ and $\mathrm{Var}(Z(c)) = x_R(c) + x_L(c)$. Since

$x_L(c)/x_R(c) \longrightarrow q^2 < 1$ and $x_R(c) \longrightarrow \infty$, we have $[x_R(c) - x_L(c)]/\sqrt{x_R(c) + x_L(c)} \longrightarrow \infty$. Chebyshev's inequality yields

$$\Pr\{L \text{ wins}\} \leq \Pr\{Z(c) \leq 0\} \leq \frac{x_R(c) + x_L(c)}{[x_R(c) - x_L(c)]^2} \longrightarrow 0.$$

Given $\delta > 0$, choose $c(\delta) \in (0, \bar{c})$ sufficiently small that, for every $c \in (0, c(\delta))$, independent abstention is strictly optimal and the preceding probability is below δ. For each such c, let $\bar{N}(c) = 1 + \lceil \max\{mx_R(c)/\alpha_R, mx_L(c)/\alpha_L\} \rceil$. For every $N \geq \bar{N}(c)$, set $\beta_R^N(c) = mx_R(c)/(\alpha_R N)$, $\beta_L^N(c) = mx_L(c)/(\alpha_L N)$, and $\sigma_N = 0$. Both mixing probabilities lie in $(0, 1)$. Poisson thinning reproduces the exact intensities $x_R(c)$ and $x_L(c)$, so both organized types remain indifferent. These strategies therefore form an equilibrium for every $N \geq \bar{N}(c)$, and candidate L wins with probability lower than δ. ■

I.5 Proof of Proposition 4

Proposition 4: There exist $\vartheta^* > 0$ and $\bar{N} < \infty$ such that, whenever $\vartheta < \vartheta^*$, every equilibrium satisfies $\sigma_N = 0$ for every $N \geq \bar{N}$.

Proof: Fix $\vartheta_0 > 0$ and restrict attention to $\vartheta \leq \vartheta_0$, recalling that $v_I > 0$. Full participation by the independents cannot be an equilibrium for all sufficiently large N. Indeed, since $\alpha_I > \alpha_R$, if $\sigma_N = 1$ then L wins with probability approaching one uniformly over the group strategies, so the independent voter's pivotal probability approaches zero, whereas $c/v_I \geq c/(v_L + \vartheta_0) > 0$.

Suppose, toward a contradiction, that $\sigma_N \in (0, 1)$. Write $T_I^N = \sigma_N \alpha_I N$, $T_R^N = \alpha_R \beta_R^N N/m$, and $T_L^N = \alpha_L \beta_L^N N/m$. Independent indifference gives $\mathcal{P}_I = c/v_I \geq c/(v_L + \vartheta_0)$. The online appendix proves that this lower bound implies a finite uniform upper bound $\overline{M}$ on T_I^N, T_R^N, T_L^N. Consequently, for all sufficiently large N, $\beta_R^N, \beta_L^N < 1$. It also proves that there exist $\vartheta_1 > 0$, $\bar{N}_1 < \infty$, and $\underline{M} > 0$ such that $T_R^N \geq \underline{M}$ in every such equilibrium whenever $\vartheta \in [0, \vartheta_1]$ and $N \geq \bar{N}_1$.

Let $S_i(T_R^N, T_L^N)$ denote the probability that the difference between the numbers of voting R-groups and voting L-groups is i. The independent-voter indifference condition is:

$$\frac{c}{v_I} = \mathcal{P}_I(\sigma_N, \beta) = \frac{1}{2}\left[\begin{array}{c} S_0\left(\frac{\alpha_R \beta_R^N N}{m}, \frac{\alpha_L \beta_L^N N}{m}\right) e^{-\sigma_N \alpha_I N} \\ + \sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R \beta_R^N N}{m}, \frac{\alpha_L \beta_L^N N}{m}\right) \cdot \left(\frac{(\sigma_N \alpha_I N)^{mi-1} e^{-\sigma_N \alpha_I N}}{(mi-1)!} + \frac{(\sigma_N \alpha_I N)^{mi} e^{-\sigma_N \alpha_I N}}{(mi)!}\right) \end{array}\right] \tag{28}$$

An independent is pivotal if the groups are tied, or if R is k groups ahead and either $mk - 1$ or

mk other independents vote for L. We also have that:

$$\begin{aligned}
\frac{c}{v_L} &\geq \mathcal{P}_L\left(\sigma_N, \beta\right) \\
&\geq \frac{1}{2}\left[\begin{array}{c} S_0\left(\frac{\alpha_R\beta_R^N N}{m}, \frac{\alpha_L\beta_L^N N}{m}\right) e^{-\sigma_N\alpha_I N} \\ +\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^N N}{m}, \frac{\alpha_L\beta_L^N N}{m}\right)\left(\begin{array}{c} \frac{(\sigma_N\alpha_I N)^{mi-1}e^{-\sigma_N\alpha_I N}}{(mi-1)!} + \frac{(\sigma_N\alpha_I N)^{mi}e^{-\sigma_N\alpha_I N}}{(mi)!} \\ +\sum_{j=2}^{m}\frac{(\sigma_N\alpha_I N)^{(mi-j)}e^{-\sigma_N\alpha_I N}}{(mi-j)!} \end{array}\right) \end{array}\right] \\
&= \mathcal{P}_I\left(\sigma_N, \beta\right) + \frac{1}{2}\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^N N}{m}, \frac{\alpha_L\beta_L^N N}{m}\right)\cdot\sum_{j=2}^{m}\frac{\left(\sigma_N\alpha_I N\right)^{(mi-j)}e^{-\sigma_N\alpha_I N}}{(mi-j)!} \\
\Longrightarrow \frac{c}{v_L} - \frac{c}{v_I} &\geq \frac{1}{2}\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^N N}{m}, \frac{\alpha_L\beta_L^N N}{m}\right)\cdot\sum_{j=2}^{m}\frac{\left(\sigma_N\alpha_I N\right)^{(mi-j)}e^{-\sigma_N\alpha_I N}}{(mi-j)!} = \chi^N \qquad (29)
\end{aligned}$$

The first inequality uses $\beta_L^N < 1$. The second follows because an L-group is pivotal in all the events that make an independent pivotal and also in the events retained in χ^N. In particular,

$$\chi^N \geq \frac{1}{2}S_1(T_R^N, T_L^N)e^{-T_I^N}.$$

If $\vartheta < 0$, the left-hand side of (29) is negative while $\chi^N \geq 0$, an immediate contradiction. It therefore remains to consider $\vartheta \in [0, \min\{\vartheta_0, \vartheta_1\}]$. On this range, for $N \geq \bar{N}_1$, the uniform bounds give $T_R^N \in [\underline{M}, \overline{M}]$, $T_L^N \in [0, \overline{M}]$, and $T_I^N \leq \overline{M}$. Continuity and positivity of the Skellam probability imply

$$\underline{S} := \min_{x\in[\underline{M},\overline{M}],\, y\in[0,\overline{M}]} S_1(x,y) > 0, \qquad \chi^N \geq \frac{1}{2}\underline{S}e^{-\overline{M}}.$$

On the other hand, $c/v_L - c/v_I = c\vartheta/[v_L(v_L+\vartheta)]$. Choose $\vartheta^* \leq \min\{\vartheta_0, \vartheta_1\}$ so that $c\vartheta^*/[v_L(v_L + \vartheta^*)] < \underline{S}e^{-\overline{M}}/2$. Then (29) is impossible for every $\vartheta < \vartheta^*$ and all sufficiently large N. Hence no equilibrium can have $\sigma_N \in (0,1)$. Since $\sigma_N = 1$ has already been excluded, every equilibrium has $\sigma_N = 0$. Taking $\bar{N}$ to be the largest of the finite thresholds used above proves the proposition. ■

I.6 Proof of Proposition 5

If voters use independent strategies (i.e. $m = 1$), total turnout is Poisson with mean $\mu \equiv \mu_R + \mu_L$ and variance $\sigma_T^2 = \mu$. Hence $\sigma_T/\mu = 1/\sqrt{\mu} \to 0$. Similarly, the margin of victory is the difference of two independent Poisson random variables, so $\sigma_{MV}^2 = \mu_R + \mu_L = \mu$ and therefore $\sigma_{MV}/\mu = 1/\sqrt{\mu} \to 0$ as $\mu \to \infty$. ■

II Appendix B: Empirical Extensions and Robustness

This appendix collects the empirical extensions and robustness exercises referenced in Sections 4.2 and 5. Section II.1 reports the grouped-model estimates within each electorate-size stratum. Section II.2 reports the coefficient estimates of the $m = 1$ benchmark. Section II.3 re-estimates the full model separately by historical period. Section II.4 re-estimates the model with an extended covariate set and with the dataset restricted to the first referendum in a jurisdiction. Section II.5 examines fixed noise rules and the sensitivity of the estimates to alternative samples and covariates. Section II.6 reports additional diagnostics for the alternative voting models of Section 5, and Section II.7 validates the estimated models out of sample, using ten-fold cross-validation that holds out entire counties. Throughout, the main estimation sample consists of the 1,724 previously-dry Texas local-option referenda with at least 100 eligible voters, and the size strata are those of the main text: small jurisdictions have $N < 1{,}500$, medium jurisdictions $1{,}500 \leq N < 5{,}000$, and large jurisdictions $N \geq 5{,}000$. The numerical procedures used here and throughout the paper—the likelihood implementation, fixed-noise sensitivity analysis, county-block bootstrap, and common scoring of the alternative models—are documented in detail in the Online Appendix (Appendix III.1).[42] All tables for this appendix are collected in Section II.8 and all figures in Section II.9.

II.1 Grouped-model estimates by electorate-size stratum

Table 9 reports the maximum-likelihood estimates of the grouped-vote model separately for the small, medium, and large electorate strata, complementing the full-sample estimates of Table 7 in the main text. All strata use the prespecified mean-zero noise rule with Poisson component mean $0.005N_i$. Standard errors use 200 county-block bootstrap draws, holding that rule fixed. The table reports coefficients for the two cost/value equations and the log group-share equation, $\log(m/N)$.

II.2 Estimates for the $m = 1$ benchmark

Table 10 reports the maximum-likelihood estimates of the $m = 1$ benchmark of Section 5, on the full sample and separately within each electorate-size stratum, using the same covariates as the two cost/value equations of the grouped model. All fits use the prespecified Poisson component mean $0.005N_i$. Standard errors come from the 200-draw county-block bootstrap described in the Online Appendix.

[42] For computational speed, we use the Myerson approximations (Equations (1) and (2)). Re-estimating the main full-sample specification with exact Poisson pivotal probabilities leaves all fitted integer group sizes unchanged and changes predicted turnout by an average absolute amount of 0.033 percentage points. This check concerns the main specification.

II.3 Stability across time

The sample spans eight decades, from 1937 through 2019. Figure 7 shows how the referenda are distributed over that period and how their outcomes evolved: local-option elections occur throughout, with the heaviest activity in the 1970s, 2000s, and 2010s, and the share won by the "for" side rising over time.

Pooling eight decades raises the concern that the estimates average over structurally different eras: Texas urbanized, communication technologies changed beyond recognition, and the organizations that coordinate civic life evolved. To address this concern, we split the sample into elections held before 1970 (494 elections), in 1970–1989 (576 elections), and in 1990 or later (654 elections), and re-estimate the grouped and $m = 1$ models separately in each period. Each period uses the same prespecified Poisson component mean $0.005N_i$, with coefficients re-estimated on that period's elections. County-block bootstrap standard errors use 200 draws and hold the noise rule fixed. Table 11 reports the grouped estimates and standard errors. The grouped model improves on the $m = 1$ benchmark by 153,344, 125,699, and 118,092 log-likelihood points in the three periods, respectively. Results on the size of groups and tails are qualitatively similar to the main specification.

Figure 8 complements the table with a descriptive plot of fitted group size by decade and electorate-size stratum, computed from the main full-sample estimates. Movements in the pooled mean $\hat{m}$ across decades largely reflect the changing composition of jurisdictions holding referenda, rather than changes in the group-size equation itself.

II.4 Extended covariates and repeated referenda

The baseline covariates were chosen because they are observed for the full historical sample and match the covariates used by Coate et al. (2008). The extended specification adds seven decade indicators (the pre-1950 period is omitted), an indicator for elections held on a statewide general, constitutional-amendment, or primary election date, standardized distance from the county centroid to the nearest Texas postsecondary institution, and the standardized county population shares aged 18–24 and 65 or older. The two cost/value equations include all of these variables; the group-share equation adds the decade, distance, and age variables to its baseline covariates.

We estimate the extended model on the full sample (1,724 elections) and on first referenda (1,060 elections). The latter sample retains qualifying propositions held on each jurisdiction's earliest recorded election date, including multiple propositions on that date. Since a second referendum only occurs in a jurisdiction if the first fails, we ensured that the results did not meaningfully change because of that selection.

Table 12 reports every coefficient in every equation, with 200-draw county-block bootstrap

standard errors for both the $m = 1$ and grouped specifications. The fit statistics at the bottom of the table show that the grouped model improves on the $m = 1$ benchmark by 426,079 log-likelihood points in the full extended sample and 223,307 in the first-referendum extended sample. The group sizes and tail probabilities are again qualitatively similar to the main specification.

II.5 Noise and specification robustness

Table 13 compares two mean-zero noise specifications for the error term in (18). Panel A uses the main specification, Skellam noise with mean zero made up of two independent Poissons with means μ_i. As mentioned, the advantage of using a symmetric Skellam distribution is that it yields an integer-valued, bell-shaped, mean-zero noise term that is approximately normally distributed for large N_i. The rows in Table 13 have different values of μ_i, the first five vary the noise proportionally with the electorate, the last row uses a fixed $\mu_i = 10$. Panel B uses a single independent Poisson with mean μ_i, shifted leftward so that it has mean zero. The five rows similarly vary μ_i proportionally with the electorate.

The grouped model has a higher likelihood in every row, by at least 400,000 log-likelihood units. At the baseline Skellam mean $0.005N_i$ and the corresponding shifted-Poisson mean $0.01N_i$, both give mean group size 42.26, with standard errors 3.05 and 3.37, respectively. Mean fitted group size ranges from 37.71 to 49.34 across the proportional Skellam specifications and from 37.52 to 51.86 under shifted-Poisson noise. Mean turnout is 29.05% and 28.92%, respectively.

Note that the estimated group size and likelihood are strictly increasing in the noise parameter, for all models. The larger the means of the Poissons that make up the Skellam (or the independent shifted Poisson), the higher the variance of the results mechanically. We chose a reasonable intermediate value, which ensures that the noise term does not dominate the likelihood. The main results we report are not particularly sensitive to the noise parameter.

Table 14 compares fourteen sample and covariate specifications. Each grouped/$m = 1$ pair uses the same observations, common covariates in the cost/value equations, and Poisson component mean $0.005N_i$. The table reports within-row likelihood gains and mean group sizes with 200-draw county-block bootstrap standard errors.

The first three specifications replace the baseline electorate floor of 100 with 50, 200, and 500 eligible voters. The next two use the extended covariates on the full and first-referendum samples described above. The CCM specification uses the 141 elections in the original replication data with $0 < N \leq 900$, without imposing our baseline floor of 100. It retains the original two proposition indicators and omits congregation size, which is unavailable in those data. The following row restricts the baseline specification to first referenda.

Four specifications examine the treatment of justice precincts. A justice precinct is typically a split of a county into equally-sized precincts, but our data are not granular enough to ensure that we always know the population of the precinct in question. We report the results of several alternative definitions.

The *County/city only* specification retains only county and city elections (1,178 elections). *Exclude all justice precincts* excludes all justice-precinct elections while retaining four other local-unit elections (1,182 elections). *Exclude uncertain precincts* retains the county and city elections and adds only precinct elections whose population denominators use high- or medium-confidence precinct counts (1,236 elections). All three retain the other baseline eligibility requirements. Mean fitted group sizes are 36.31, 36.27, and 37.66, respectively, and tails and likelihoods are qualitatively similar.

The restaurant specification retains all 1,724 elections, pools restaurant-only mixed-beverage propositions with the bars category, and omits the separate restaurant indicator, matching the definitions from Coate et al. (2008) directly. The fourth precinct specification retains the baseline sample and adds 11 eligible historical or combined precinct elections, yielding 1,735 elections. Its mean fitted group size is 42.77. The Melissa specification changes only the audited Melissa proposition from the omitted category to the all-liquor off-premise category.[43] The sample remains at 1,724, and mean fitted group size is 42.26.

Finally, we exclude standard November election dates and regular first-primary dates. This calendar-based restriction removes 342 elections (333 on standard November dates and nine on regular first-primary dates), leaving 1,382 elections. Mean fitted group size is 41.26, with a county-block bootstrap standard error of 2.80, compared with 42.26 in the main sample. The grouped model improves on the $m = 1$ benchmark by 334,401 log-likelihood points.

Overall, the results are qualitatively similar and not sensitive to these various changes in sample definition. In particular, we reject the null hypothesis that groups do not improve the fit of the model in each case at any conventional significance level.

II.6 Additional diagnostics for the alternative voting models

Two further diagnostics complement the model comparison of Section 5.

First, although Table 8 shows that the median-shock finite-count GRU adaptation generates substantial dispersion in the margin of victory, its implied turnout distribution is too concentrated.

[43]The TABC report for the May 9, 2015 election in the City of Melissa (in Collin County, Texas) is contradictory – the proposition itself is described as "the legal sale of beer and wine for off-premises consumption" but the concluding sentence says that the election permitted "all alcoholic beverages for off-premises consumption." We use the former in the main specification, and the latter for this specification.

Figure 9 compares the empirical turnout histogram with the model's simulated density. The simulated density has a pronounced peak near 0.17 and a smaller secondary peak around 0.20. Observed turnout is much more dispersed, with substantially more mass at very low turnout and a longer right tail. The model's turnout standard deviation is 0.097, compared with 0.181 in the data. Thus, in this benchmark, integrating over support heterogeneity generates considerable variation in electoral margins without reproducing the observed variation in total participation. This diagnostic concerns the fitted adaptation conditional on the common benefit/cost shock being at its median.

Second, Figure 10 repeats the election-level diagnostic of Figure 5 for the expressive-voting model. The figure is constructed as follows. Nine elections are selected at fixed quantiles of the electorate-size distribution, from the 5th to the 99th percentile, so that the panels span the sample without discretion in the choice of elections. For each election, we simulate 300 outcomes from the fitted expressive model and 300 from the fitted grouped model, plot the simulated (turnout, margin) pairs, and mark the observed outcome with a star. Because expressive participation responds neither to the electorate size nor to any equilibrium object, the model's predictive distribution is a tight cluster whose spread shrinks as N grows and whose location frequently misses the realization in both dimensions. The grouped model's predictive distribution remains two-dimensional at every electorate size and typically covers the observed outcome.

II.7 Out-of-sample validation

The in-sample comparisons of Section 5 could in principle reflect overfitting: the grouped model uses six additional free parameters. To address this concern, we validate both models out of sample using ten-fold cross-validation by county. We assign the 216 counties to ten disjoint folds, hold out one fold at a time, re-estimate both the $m = 1$ benchmark and the grouped model on the remaining counties, and score the held-out elections. Each of the 1,724 elections is held out exactly once; the folds contain 172 or 173 elections. Both models retain the prespecified Poisson component mean $0.005N_i$ in every fold. Standardization, starting values, and coefficient estimation use only the training counties, so the held-out outcomes play no role in estimation.

Panel A of Table 15 reports the election-level prediction fit, pooling the held-out elections across the ten folds. On held-out turnout the grouped model attains a higher correlation (0.536 against 0.484 for $m = 1$) and lower prediction errors (mean absolute error 0.120 against 0.126). Figure 11 compares fitted and realized turnout election by election, using the full-sample fit (left panel) and the pooled county-held-out predictions (right panel).

The sharp out-of-sample distinction between the models concerns the margin of victory. Panel B of Table 15 reports the dispersion and tail probabilities of the margin of victory on the held-out

elections, comparing the empirical distribution with the predictive distributions at each election's training-fold estimates. The grouped model captures substantially more of the out-of-sample dispersion, although it still understates the observed tails: the standard deviation of the margin is 0.102 in the held-out data and 0.072 in the grouped model, against 0.023 under $m = 1$; margins of victory larger than 10 percent occur in 24.7% of the held-out data, compared with 13.9% under the grouped model and 0.6% under $m = 1$. The same ordering holds at every reported threshold. Figure 12 makes the comparison visually: for each electorate-size stratum it plots the empirical density of the absolute margin of victory on the held-out elections against the simulated densities from the two models. The $m = 1$ benchmark concentrates most of its mass near a zero margin, while the grouped model generates greater dispersion. The failure of the pivotal-voter model in the margin dimension thus persists out of sample, and introducing organized groups reduces it.

To examine variation across held-out county groups, Panel C of Table 15 reports the unweighted mean and range of each fit statistic across the ten folds, and Panel D reports the corresponding summaries of the margin-of-victory tails. These are disjoint county folds from one cross-validation exercise, rather than repeated random splits. The grouped model attains a higher mean turnout correlation (0.531 against 0.480) and lower turnout errors (mean absolute error 0.120 against 0.126) than the $m = 1$ benchmark, with the same turnout ordering in seven of ten folds. At the 0.10 margin threshold, the tail ordering is stable across all ten folds: the grouped-model entries range over $[0.094, 0.180]$ and the $m = 1$ entries over $[0.004, 0.011]$, against empirical values in $[0.185, 0.308]$.

II.8 Tables

Table 9: Grouped-Vote Estimates by Electorate Size

Covariate	Small	Medium	Large
Panel A: $\log \gamma_R$			
Intercept	-0.812^{***}	-0.893	-1.452
	(0.288)	(0.854)	(1.093)
$\log(N)$	-0.245^{***}	-0.168	-0.169^{**}
	(0.040)	(0.109)	(0.084)
Baptist share	-0.308^{*}	-0.662	-0.212
	(0.182)	(0.480)	(1.578)
City	-0.001	-0.227^{*}	0.090
	(0.087)	(0.120)	(0.210)
Weekend	0.022	0.059	0.269^{***}
	(0.034)	(0.063)	(0.098)
All-alcohol/non-mixed	-0.059	-0.347^{**}	-0.388^{*}
	(0.061)	(0.144)	(0.230)
Mixed-beverage unrestricted	0.071	-0.039	-0.308^{*}
	(0.054)	(0.101)	(0.165)
Restaurant-only mixed-beverage	-0.476	-0.254	0.061
	(0.432)	(0.226)	(0.414)
Panel B: $\log \gamma_L$			
Intercept	-0.965^{***}	-0.715	-1.068
	(0.306)	(0.786)	(1.043)
$\log(N)$	-0.206^{***}	-0.183^{*}	-0.192^{**}
	(0.043)	(0.100)	(0.080)
Baptist share	-0.302^{*}	-0.770^{**}	-0.341
	(0.179)	(0.365)	(1.434)
City	-0.076	-0.294^{**}	-0.044
	(0.073)	(0.119)	(0.162)
Weekend	0.043	0.048	0.244^{***}
	(0.033)	(0.051)	(0.084)
All-alcohol/non-mixed	-0.083^{**}	-0.194^{**}	-0.284
	(0.035)	(0.082)	(0.187)
Mixed-beverage unrestricted	-0.035	0.025	-0.225^{*}
	(0.065)	(0.058)	(0.132)
Restaurant-only mixed-beverage	-0.796	-0.383	0.027
	(0.545)	(0.321)	(0.482)
Panel C: $\log(m/N)$			
Intercept	2.623	9.244	0.341
	(2.824)	(16.676)	(4.309)
$\log(N)$	-1.443	-2.988	-0.717
	(0.925)	(4.244)	(0.918)
$[\log(N)]^2$	0.065	0.160	0.015
	(0.075)	(0.270)	(0.049)
Baptist share	0.113	0.131	0.277
	(0.213)	(0.435)	(0.383)
City	-0.081	-0.338^{*}	-0.320^{***}
	(0.069)	(0.177)	(0.049)
SBC congregation size (std.)	0.010	0.000	0.025
	(0.040)	(0.029)	(0.038)
Elections	640	561	523
Log likelihood	$-7{,}685.1$	$-9{,}066.6$	$-12{,}139.6$
Mean $\hat{m}$	11.54	31.02	83.70
Median $\hat{m}$	11.00	30.00	70.00

Notes: The native equations are $\log \gamma_s = X^\top \xi_s$ for $s \in \{R, L\}$. Before rounding, $\log(m/N) = X_m^\top \xi_m$. County-block bootstrap standard errors (200 paired draws) are in parentheses. Coefficient stars use normal approximations: *, **, and *** indicate the 10, 5, and 1 percent levels. Small, medium, and large electorates satisfy $N < 1{,}500$, $1{,}500 \le N < 5{,}000$, and $N \ge 5{,}000$, respectively.

Table 10: Independent-Voter Estimates

Covariate	Full	Small	Medium	Large
Panel A: $\log \gamma_R$				
Intercept	−1.145**	−1.140	−2.965	−8.017
	(0.468)	(0.971)	(6.071)	(16.205)
$\log(N)$	−0.441***	−0.373***	−0.111	0.060
	(0.034)	(0.113)	(0.718)	(1.126)
Baptist share	0.684	−0.396	−2.891	4.905
	(0.650)	(0.491)	(2.125)	(11.061)
City	−0.113	−0.110	−0.256	−0.185
	(0.099)	(0.508)	(0.604)	(1.040)
Weekend	0.068	0.095	−0.147	−0.024
	(0.068)	(0.176)	(0.461)	(0.536)
All-alcohol/non-mixed	0.021	−0.101	0.093	0.333
	(0.129)	(0.328)	(1.164)	(2.852)
Mixed-beverage unrestricted	0.008	−0.079	0.442	0.229
	(0.088)	(0.298)	(0.833)	(2.974)
Restaurant-only mixed-beverage	0.076	−0.068	0.328	0.211
	(0.153)	(2.379)	(3.624)	(1.186)
Panel B: $\log \gamma_L$				
Intercept	−1.040**	−1.266	−2.715	−7.814
	(0.441)	(0.894)	(5.925)	(16.054)
$\log(N)$	−0.449***	−0.353***	−0.140	0.045
	(0.033)	(0.118)	(0.703)	(1.115)
Baptist share	0.610	−0.402	−2.842	4.807
	(0.634)	(0.475)	(2.095)	(11.007)
City	−0.113	−0.145	−0.253	−0.183
	(0.097)	(0.496)	(0.614)	(1.037)
Weekend	0.072	0.104	−0.138	−0.019
	(0.067)	(0.165)	(0.449)	(0.534)
All-alcohol/non-mixed	0.017	−0.100	0.086	0.327
	(0.125)	(0.310)	(1.147)	(2.844)
Mixed-beverage unrestricted	0.003	−0.095	0.423	0.223
	(0.087)	(0.291)	(0.820)	(2.964)
Restaurant-only mixed-beverage	0.077	−0.081	0.316	0.211
	(0.150)	(2.431)	(3.662)	(1.182)
Elections	1,724	640	561	523
Log likelihood	−487,406.8	−25,575.9	−78,072.5	−366,035.1
Mean $\hat{m}$	1.00	1.00	1.00	1.00
Median $\hat{m}$	1.00	1.00	1.00	1.00

Notes: The native equations are $\log \gamma_s = X^\top \xi_s$ for $s \in \{R, L\}$. The model imposes $m = 1$. County-block bootstrap standard errors (200 paired draws) are in parentheses. Coefficient stars use normal approximations: *, **, and *** indicate the 10, 5, and 1 percent levels. Small, medium, and large electorates satisfy $N < 1{,}500$, $1{,}500 \leq N < 5{,}000$, and $N \geq 5{,}000$, respectively.

Table 11: Grouped-Vote Estimates by Time Period

Covariate	Before 1970	1970–1989	1990 onward
Panel A: $\log \gamma_R$			
Intercept	-2.073^{**}	-0.282	-1.308^{***}
	(0.918)	(0.248)	(0.384)
$\log(N)$	-0.065	-0.208^{***}	-0.231^{***}
	(0.110)	(0.029)	(0.025)
Baptist share	-2.750	-2.196^{***}	0.764
	(2.159)	(0.696)	(0.660)
City	0.301	-0.315^{***}	0.174
	(0.453)	(0.075)	(0.243)
Weekend	0.273	0.015	0.003
	(0.364)	(0.061)	(0.131)
All-alcohol/non-mixed	-0.444	-0.025	-0.177
	(0.435)	(0.074)	(0.143)
Mixed-beverage unrestricted	-0.063	0.071	-0.182
	(0.215)	(0.120)	(0.177)
Restaurant-only mixed-beverage	–	–	-0.056
			(0.127)
Panel B: $\log \gamma_L$			
Intercept	-1.401^{*}	-0.413	-1.921^{***}
	(0.769)	(0.253)	(0.480)
$\log(N)$	-0.123	-0.198^{***}	-0.192^{***}
	(0.093)	(0.025)	(0.030)
Baptist share	-2.503	-1.555^{***}	1.098
	(1.815)	(0.550)	(0.857)
City	0.134	-0.397^{***}	0.274
	(0.366)	(0.065)	(0.300)
Weekend	0.245	0.034	-0.080
	(0.311)	(0.054)	(0.167)
All-alcohol/non-mixed	-0.367	-0.063	-0.192
	(0.359)	(0.051)	(0.190)
Mixed-beverage unrestricted	-0.063	0.012	-0.236
	(0.179)	(0.083)	(0.235)
Restaurant-only mixed-beverage	–	–	-0.084
			(0.147)
Panel C: $\log(m/N)$			
Intercept	0.333	0.743	-1.773^{***}
	(1.281)	(0.714)	(0.420)
$\log(N)$	-0.717^{**}	-0.754^{***}	-0.085
	(0.357)	(0.182)	(0.106)
$[\log(N)]^2$	0.015	0.015	-0.029^{***}
	(0.024)	(0.012)	(0.007)
Baptist share	0.277	-0.199	0.165
	(0.215)	(0.214)	(0.103)
City	-0.320^{***}	-0.411^{***}	-0.308^{***}
	(0.120)	(0.068)	(0.040)
SBC congregation size (std.)	0.023	0.037	0.040^{*}
	(0.047)	(0.049)	(0.022)
Elections	494	576	654
Log likelihood	$-8{,}837.5$	$-9{,}420.9$	$-10{,}073.5$
Mean $\hat{m}$	49.90	34.09	35.38
Median $\hat{m}$	37.00	26.00	27.00

Notes: The native equations are $\log \gamma_s = X^{\top}\xi_s$ for $s \in \{R, L\}$. Before rounding, $\log(m/N) = X_m^{\top}\xi_m$. County-block bootstrap standard errors (200 paired draws) are in parentheses. Coefficient stars use normal approximations: *, **, and *** indicate the 10, 5, and 1 percent levels. – denotes a coefficient not separately estimable.

Table 12: Extended-Covariate Estimates

Covariate	Full $m = 1$	Full grouped	First-ref. $m = 1$	First-ref. grouped
Panel A: $\log \gamma_R$				
Intercept	−0.742	−0.267	−0.149	−0.596**
	(0.514)	(0.314)	(0.454)	(0.304)
log(N)	−0.459***	−0.268***	−0.536***	−0.273***
	(0.034)	(0.024)	(0.046)	(0.023)
Baptist share	−0.004	−0.348*	0.462	−0.370*
	(0.599)	(0.201)	(0.534)	(0.208)
City	−0.217*	−0.452***	−0.061	−0.426***
	(0.114)	(0.080)	(0.160)	(0.069)
Weekend	0.041	0.030	0.055	0.049
	(0.092)	(0.054)	(0.131)	(0.111)
All-alcohol/non-mixed	−0.130	−0.051*	−0.153	−0.119***
	(0.082)	(0.029)	(0.098)	(0.038)
Mixed-beverage unrestricted	−0.240**	0.021	−0.055	0.022
	(0.119)	(0.032)	(0.143)	(0.046)
Restaurant-only mixed-beverage	0.257**	−0.034	0.062	−0.075
	(0.119)	(0.093)	(0.143)	(0.098)
Concurrent election	−0.213*	−0.183***	0.193	−0.126
	(0.116)	(0.065)	(0.260)	(0.135)
1950s	−0.080	0.077	−0.069	−0.015
	(0.153)	(0.153)	(2.130)	(0.363)
1960s	0.227	0.060	0.059	0.230
	(0.266)	(0.155)	(0.274)	(0.269)
1970s	0.261	0.206	−0.169	0.383
	(0.229)	(0.145)	(0.222)	(0.267)
1980s	0.399	0.002	−0.331	0.344
	(0.281)	(0.209)	(0.274)	(0.279)
1990s	0.414	0.188	−0.323	0.352
	(0.379)	(0.203)	(0.419)	(0.288)
2000s	0.246	0.271	−0.174	0.573*
	(0.204)	(0.179)	(0.289)	(0.305)
2010s	0.195	0.272	−0.191	0.650**
	(0.274)	(0.188)	(0.314)	(0.300)
Distance to college (std.)	0.058	−0.064**	0.062	−0.014
	(0.050)	(0.028)	(0.063)	(0.028)
Share age 18–24 (std.)	0.027	0.001	0.024	0.017
	(0.051)	(0.027)	(0.065)	(0.035)
Share age 65+ (std.)	−0.076	−0.058***	−0.157**	−0.089***
	(0.047)	(0.022)	(0.067)	(0.026)
Panel B: $\log \gamma_L$				
Intercept	−0.656	−0.246	−0.223	−0.335
	(0.484)	(0.256)	(0.431)	(0.258)
log(N)	−0.466***	−0.261***	−0.531***	−0.278***
	(0.032)	(0.024)	(0.043)	(0.027)
Baptist share	−0.046	−0.420**	0.512	−0.428**
	(0.586)	(0.193)	(0.555)	(0.189)
City	−0.213*	−0.524***	−0.050	−0.478***
	(0.113)	(0.079)	(0.163)	(0.055)
Weekend	0.043	0.031	0.053	0.043
	(0.090)	(0.042)	(0.132)	(0.082)
All-alcohol/non-mixed	−0.128	−0.058**	−0.153	−0.119***
	(0.080)	(0.028)	(0.098)	(0.040)
Mixed-beverage unrestricted	−0.231**	−0.010	−0.051	−0.020
	(0.114)	(0.039)	(0.145)	(0.046)
Restaurant-only mixed-beverage	0.245**	−0.061	0.063	−0.098
	(0.114)	(0.114)	(0.146)	(0.118)
Concurrent election	−0.214*	−0.144**	0.208	−0.085
	(0.116)	(0.060)	(0.267)	(0.126)
1950s	−0.083	0.016	−0.085	0.072
	(0.150)	(0.117)	(2.095)	(0.296)
1960s	0.214	0.159	0.059	0.214
	(0.259)	(0.107)	(0.274)	(0.210)

Continued

Covariate	Full $m = 1$	Full grouped	First-ref. $m = 1$	First-ref. grouped
1970s	0.247	0.249**	−0.181	0.314
	(0.222)	(0.103)	(0.225)	(0.207)
1980s	0.377	0.149	−0.344	0.312
	(0.273)	(0.138)	(0.277)	(0.213)
1990s	0.386	0.318**	−0.344	0.355*
	(0.371)	(0.134)	(0.424)	(0.214)
2000s	0.243	0.161	−0.188	0.308
	(0.200)	(0.151)	(0.293)	(0.270)
2010s	0.198	0.160	−0.209	0.367
	(0.266)	(0.156)	(0.318)	(0.252)
Distance to college (std.)	0.053	−0.049*	0.063	−0.008
	(0.049)	(0.026)	(0.064)	(0.025)
Share age 18–24 (std.)	0.027	0.004	0.025	0.016
	(0.051)	(0.024)	(0.066)	(0.031)
Share age 65+ (std.)	−0.073	−0.058***	−0.161**	−0.093***
	(0.046)	(0.020)	(0.068)	(0.024)
Panel C: $\log(m/N)$				
Intercept		0.111		0.377
		(0.589)		(0.394)
$\log(N)$		−0.583***		−0.674***
		(0.138)		(0.099)
$[\log(N)]^2$		0.002		0.008
		(0.008)		(0.006)
Baptist share		0.076		0.096
		(0.157)		(0.071)
City		−0.514***		−0.378***
		(0.090)		(0.019)
SBC congregation size (std.)		0.015		−0.056***
		(0.016)		(0.009)
1950s		0.142**		0.097**
		(0.058)		(0.044)
1960s		0.262***		0.092***
		(0.076)		(0.029)
1970s		0.329***		0.057***
		(0.101)		(0.020)
1980s		0.329***		0.179***
		(0.089)		(0.025)
1990s		0.410***		0.155***
		(0.133)		(0.034)
2000s		0.241***		0.040**
		(0.090)		(0.020)
2010s		0.060		0.068***
		(0.037)		(0.020)
Distance to college (std.)		−0.077***		−0.054***
		(0.025)		(0.005)
Share age 18–24 (std.)		−0.021**		−0.004
		(0.011)		(0.005)
Share age 65+ (std.)		−0.032**		−0.078***
		(0.015)		(0.008)
Elections	1,724	1,724	1,060	1,060
Log likelihood	−455,037.3	−28,958.6	−240,680.7	−17,373.3
Mean $\hat{m}$	1.00	41.41	1.00	35.07
Median $\hat{m}$	1.00	31.00	1.00	26.00

Notes: The native equations are $\log \gamma_s = X^\top \xi_s$ for $s \in \{R, L\}$. Before rounding, $\log(m/N) = X_m^\top \xi_m$. County-block bootstrap standard errors (200 paired draws) are in parentheses. Coefficient stars use normal approximations: *, **, and *** indicate the 10, 5, and 1 percent levels. First-ref. denotes the first-referendum sample.

Table 13: Alternative Noise Specifications

Poisson mean	ℓ_1	ℓ_G	Mean $\hat{m}$ (SE)	Turnout (%)	Max. net tail (%)
Panel A: Mean-zero Skellam					
$\mu_i = 0.0025N_i$	−518,064.2	−31,910.2	37.71 (3.36)	29.14	0.004358
$\mu_i = 0.00375N_i$	−501,716.4	−30,186.9	40.25 (3.31)	29.01	0.03045
$\mu_i = 0.005N_i$	−487,406.8	−29,169.6	42.26 (3.05)	29.05	0.1075
$\mu_i = 0.0075N_i$	−462,944.3	−27,944.6	47.75 (3.11)	28.91	0.5190
$\mu_i = 0.01N_i$	−442,270.1	−27,416.6	49.34 (3.89)	29.17	1.344
$\mu_i = 10$	−542,802.6	−35,994.2	31.20 (0.97)	27.99	38.29
Panel B: Shifted Poisson					
$\mu_i = 0.005N_i$	−531,721.0	−32,070.6	37.52 (3.32)	28.94	0.03666
$\mu_i = 0.0075N_i$	−516,775.7	−30,298.0	40.06 (3.45)	28.93	0.1014
$\mu_i = 0.01N_i$	−503,028.3	−29,209.1	42.26 (3.37)	28.92	0.2365
$\mu_i = 0.015N_i$	−478,452.2	−28,101.9	45.64 (3.80)	28.83	0.7483
$\mu_i = 0.02N_i$	−456,943.0	−27,280.3	51.86 (4.56)	29.03	1.573

Notes: Both models are re-estimated on the same 1,724 elections in every row. Panel A uses the difference of two independent Poisson variables, each with mean μ_i. Panel B uses a Poisson variable with mean μ_i, shifted by an independent randomized integer to have mean zero; Appendix III.1.3 defines the shift. Panel B doubles the Poisson means in the proportional rows of Panel A to approximately match their noise variances. Parentheses give standard errors from 200 county-block bootstrap draws. The main rule is Panel A's $\mu_i = 0.005N_i$. The last column is $100 \max_i \Pr(|\varepsilon_{R,i} + \varepsilon_{L,i}| > 0.05N_i)$.

Table 14: Robustness Checks

Specification	Elections	ℓ_1	ℓ_G	$\Delta\ell$	$\Delta\ell$/Election	Mean $\hat{m}$ (SE)
Main, $N \geq 50$	1,732	−487,536.6	−29,240.0	458,296.6	264.605	42.08 (3.60)
Main, $N \geq 200$	1,658	−486,138.2	−28,502.0	457,636.2	276.017	43.68 (3.09)
Main, $N \geq 500$	1,490	−481,405.4	−26,629.5	454,776.0	305.219	47.55 (3.05)
Full sample, extended covariates	1,724	−455,037.3	−28,958.6	426,078.7	247.145	41.41 (3.17)
First referendum, extended covariates	1,060	−240,680.7	−17,373.3	223,307.4	210.667	35.07 (2.02)
CCM, $N \leq 900$	141	−3,249.2	−1,498.9	1,750.4	12.414	9.06 (0.64)
First referendum, baseline covariates	1,060	−269,143.5	−17,664.4	251,479.1	237.244	38.51 (2.44)
Legacy restaurant-only as bars mapping	1,724	−500,677.4	−29,413.6	471,263.8	273.355	40.52 (2.95)
County and city elections only	1,178	−205,846.5	−17,848.5	187,998.1	159.591	36.31 (3.87)
County/city plus high/medium-confidence precinct counts	1,236	−245,593.0	−19,104.2	226,488.9	183.243	37.66 (3.52)
Exclude all precinct-like elections	1,182	−209,075.5	−17,961.4	191,114.1	161.687	36.27 (3.38)
Alternative Melissa proposition coding	1,724	−487,480.4	−29,174.7	458,305.7	265.839	42.26 (3.00)
Include eligible historical/combined precincts	1,735	−521,671.2	−29,579.2	492,091.9	283.626	42.77 (3.41)
Exclude standard November and regular primary dates	1,382	−357,343.3	−22,942.2	334,401.1	241.969	41.26 (2.80)

Notes: ℓ_1 and ℓ_G are log likelihoods for the independent and grouped models; $\Delta\ell = \ell_G - \ell_1$. Each comparison uses the same observations within a row. County-block bootstrap standard errors (200 paired draws) are in parentheses.

Table 15: County-Held-Out Validation

Statistic	Empirical	$m = 1$	Grouped
Panel A: Pooled prediction fit			
Elections	1,724	1,724	1,724
Corr. turnout	–	0.484	0.536
MAE turnout	–	0.126	0.120
Panel B: Pooled margin-of-victory distribution			
SD margin	0.102	0.023	0.072
$\Pr(\lvert M\rvert > 0.05)$	0.485	0.049	0.392
$\Pr(\lvert M\rvert > 0.10)$	0.247	0.006	0.139
$\Pr(\lvert M\rvert > 0.15)$	0.128	0.001	0.051
$\Pr(\lvert M\rvert > 0.20)$	0.068	0.000	0.020
Panel C: Prediction fit across county folds			
Corr. turnout	–	0.480 [0.353, 0.628]	0.531 [0.421, 0.634]
MAE turnout	–	0.126 [0.107, 0.139]	0.120 [0.104, 0.132]
RMSE turnout	–	0.159 [0.138, 0.174]	0.153 [0.135, 0.169]
Panel D: Margin-of-victory distribution across county folds			
SD margin	0.101 [0.082, 0.116]	0.021 [0.018, 0.025]	0.072 [0.062, 0.083]
$\Pr(\lvert M\rvert > 0.05)$	0.486 [0.419, 0.564]	0.049 [0.031, 0.071]	0.392 [0.300, 0.432]
$\Pr(\lvert M\rvert > 0.10)$	0.247 [0.185, 0.308]	0.006 [0.004, 0.011]	0.139 [0.094, 0.180]
$\Pr(\lvert M\rvert > 0.15)$	0.128 [0.081, 0.156]	0.001 [0.000, 0.002]	0.051 [0.033, 0.077]
$\Pr(\lvert M\rvert > 0.20)$	0.068 [0.035, 0.110]	0.000 [0.000, 0.000]	0.020 [0.011, 0.035]

Notes: Ten folds hold out entire counties, predicting each election once. Panels A/B pool elections; C/D report fold means with [min, max]. Model SD is the standard deviation of the predictive mixture, including within-election variation. Tail probabilities use 1,000 simulated draws per election, weighted equally across elections. $M = \tilde{\theta}_R - \tilde{\theta}_L$.

II.9 Figures

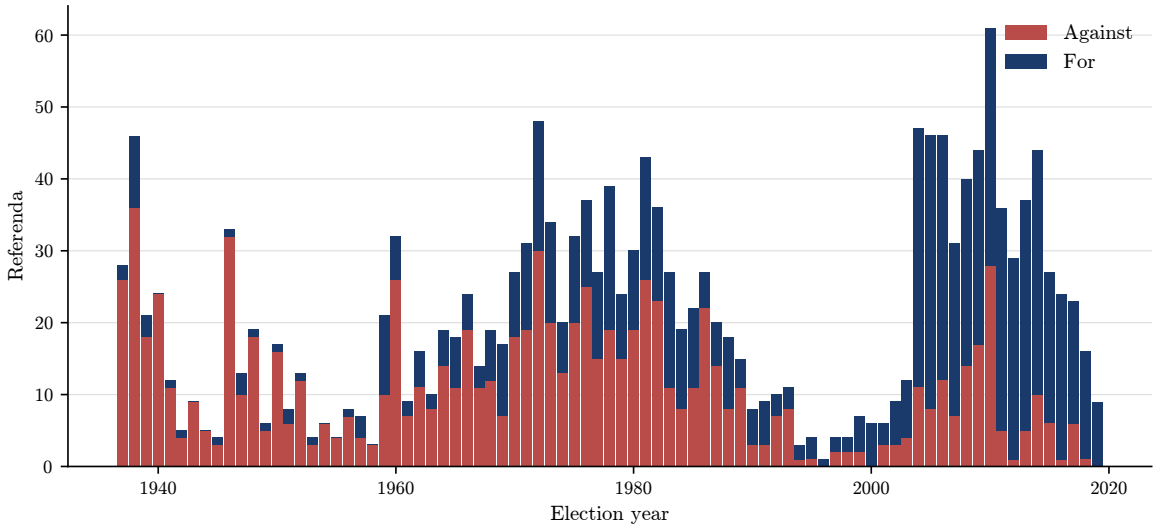


Figure 7: Liquor referenda in previously-dry jurisdictions over time. The red lower segment of each bar records elections in which "against" won (the jurisdiction stayed dry); the blue upper segment records elections in which "for" won.

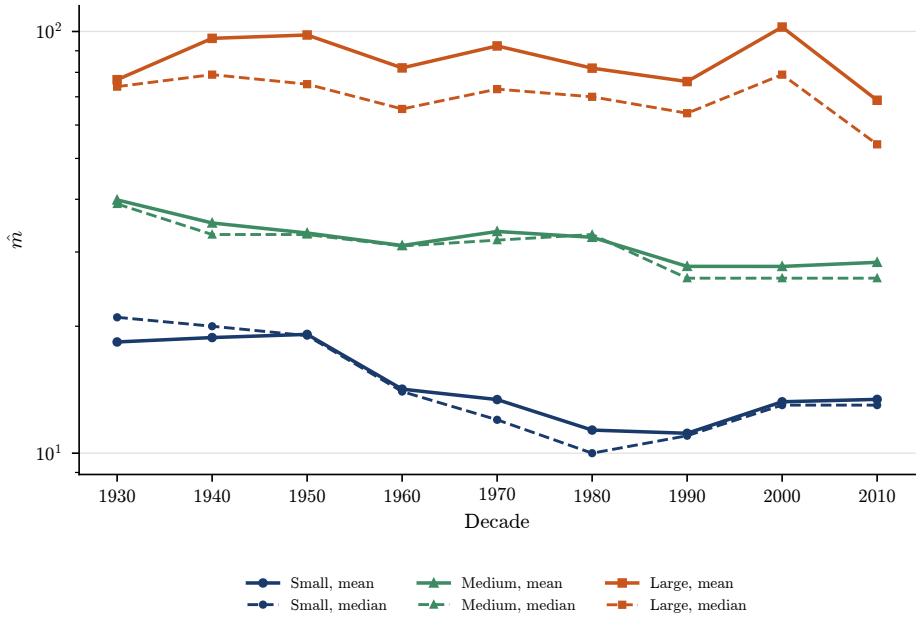


Figure 8: Fitted group size by decade and electorate-size stratum, from the main full-sample grouped model. The figure is descriptive; the period-specific structural estimates are reported in Table 11.

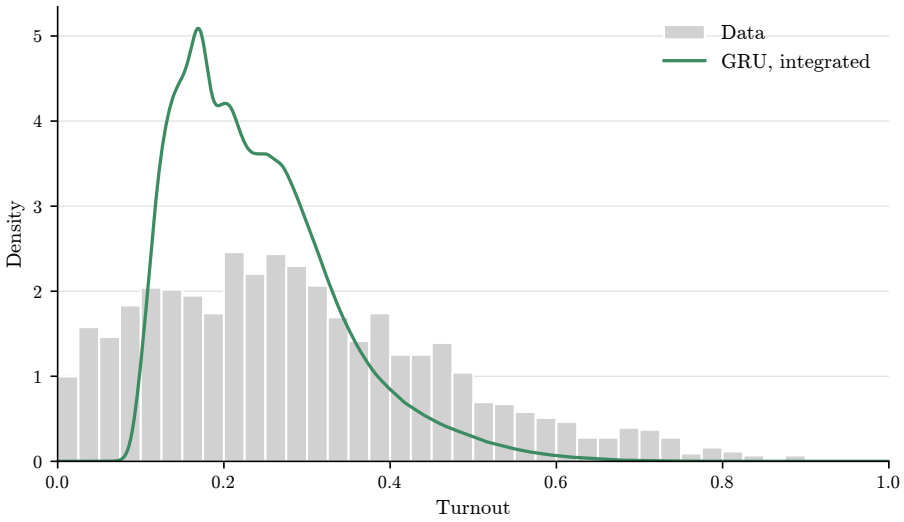


Figure 9: Simulated turnout distribution under the median-shock GRU count adaptation (green line) against the data (histogram).

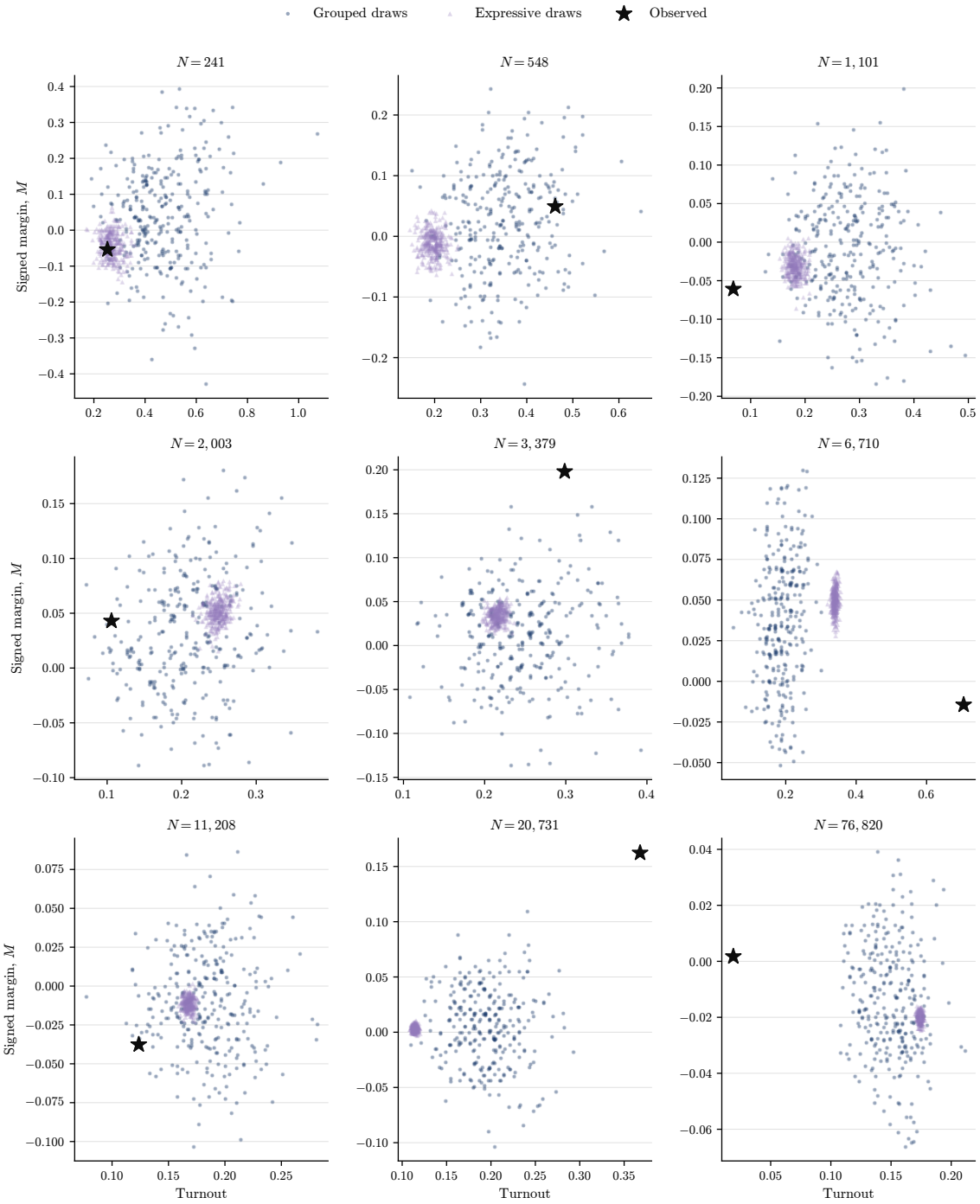


Figure 10: Election-level predictive distributions under the expressive-voting model (purple) and the grouped model (blue), for nine elections at fixed quantiles of the electorate size; the star marks the observed outcome. Each panel shows 300 simulated draws from each model.

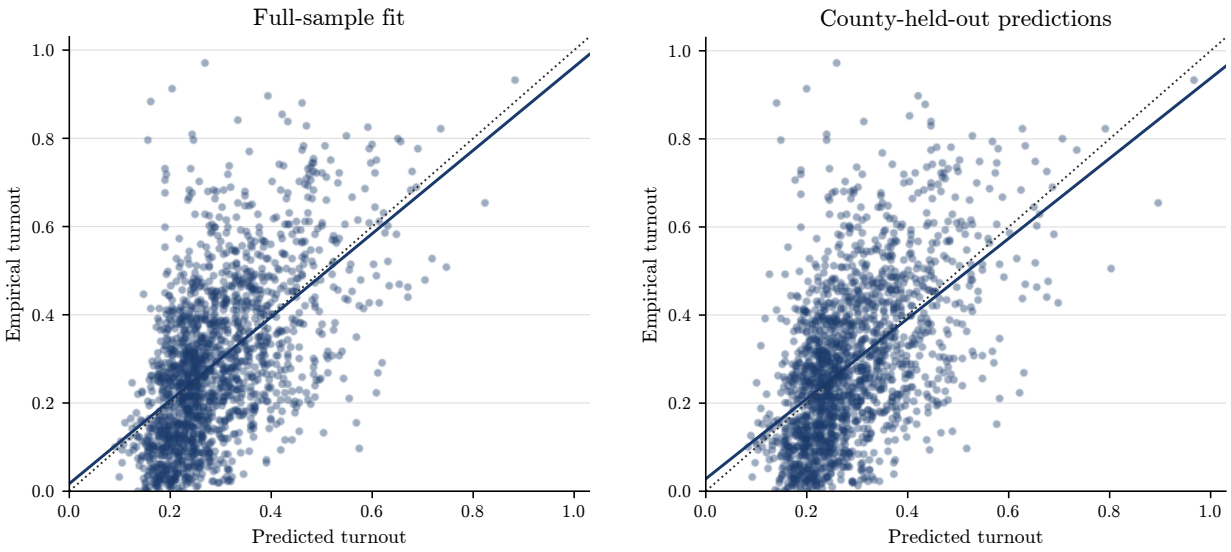


Figure 11: Fitted versus empirical turnout under the grouped model: full estimation sample (left) and held-out elections at the training-slice estimates (right). The dotted line is the 45-degree line.

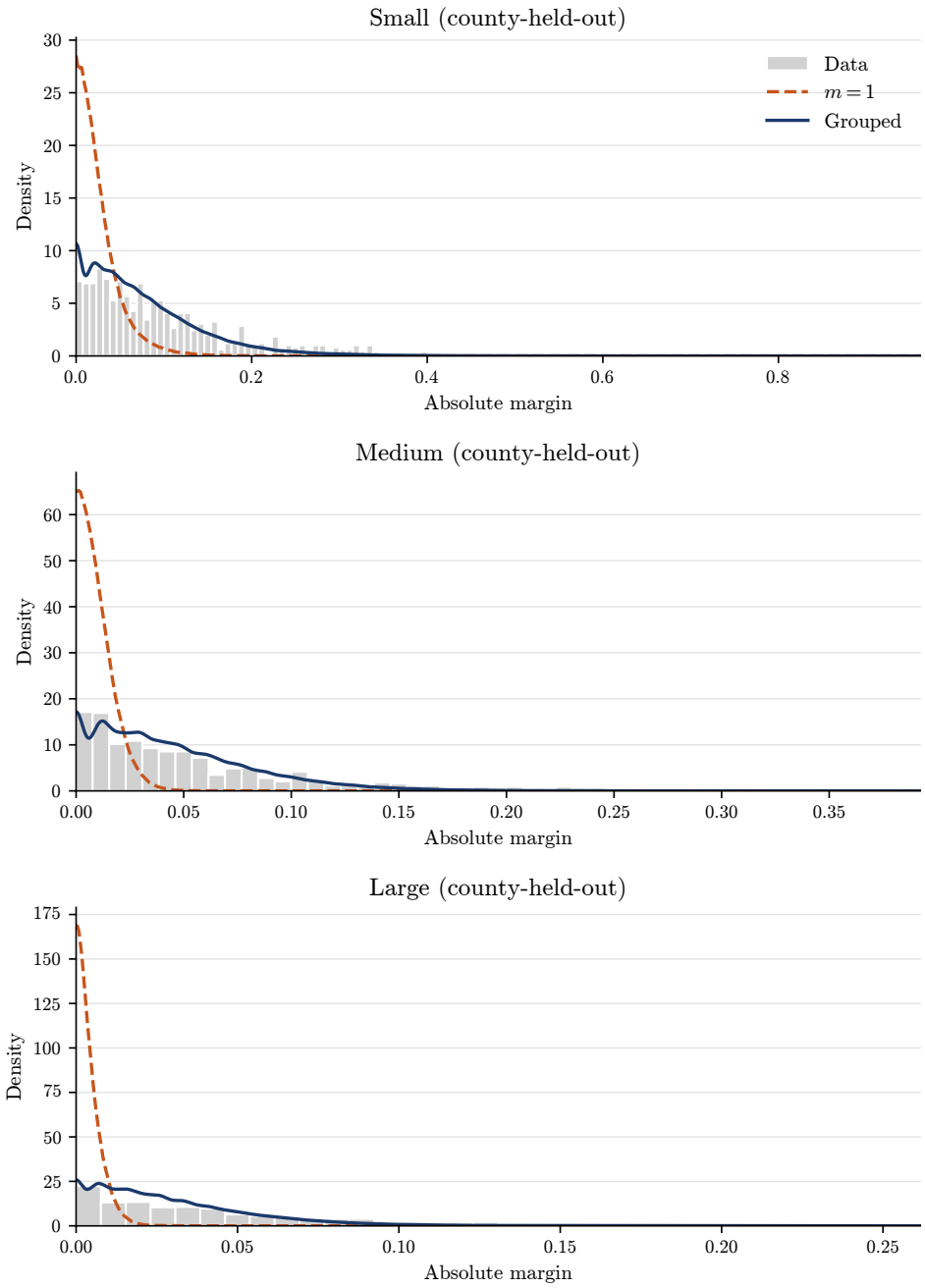


Figure 12: Absolute margin of victory in held-out counties, by electorate-size stratum. Empirical distributions are compared with simulations from the $m = 1$ and grouped models fitted to the remaining counties (Section II.7).

References

Arnold, A. and H. Sieg (2026): "The Decentralization of Liquor Policies in Texas During the Post-Prohibition Era," *International Economic Review*.

Battaglini, M., R. Morton, and T. R. Palfrey (2008): "Information Aggregation and Strategic Abstention in Large Laboratory Elections," *American Economic Review*, 98, 194–200.

——— (2010): "The Swing Voter's Curse in the Laboratory," *Review of Economic Studies*, 77, 61–89.

Berry, S. T., C. Cox, and P. A. Haile (2026): "Selective Turnout, Voting Policy, and Partisan Bias: Evidence from Multi-Level Data," .

Brennan, G. and L. Lomasky (1993): *Democracy and Decision*, Cambridge University Press, Cambridge.

Bursztyn, L., D. Cantoni, P. Funk, F. Schönenberger, and N. Yuchtman (2024): "Identifying the Effect of Election Closeness on Voter Turnout: Evidence from Swiss Referenda," *Journal of the European Economic Association*, 22, 876–914.

Coate, S. and M. Conlin (2004): "A Group Rule Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review*, 94, 1476–1504.

Coate, S., M. Conlin, and A. Moro (2008): "The Performance of Pivotal-voter Models in Small-scale Elections: Evidence from Texas Liquor Referenda," *Journal of Public Economics*, 92, 582–596.

Dellavigna, S., J. List, U. Malmendier, and G. Rao (2017): "Voting to Tell Others," *Review of Economic Studies*, 84, 143–182.

Dhillon, A. and S. Peralta (2002): "Economic Theories of Voter Turnout," *Economic Journal*, 112, F332–F352.

Downs, A. (1957): *An Economic Theory of Democracy*, New York: Harper and Row.

Feddersen, T. (2004): "Rational Choice Theory and the Paradox of Not Voting," *Journal of Economic Perspectives*, 18, 99–112.

Feddersen, T. and W. Pesendorfer (1996): "The Swing Voter's Curse," *American Economic Review*, 86, 408–424.

Feddersen, T. and A. Sandroni (2006): "A Theory of Participation in Elections," *American Economic Review*, 96, 1271–1282.

Ferejohn, J. A. and M. P. Fiorina (1974): "The Paradox of Not Voting: A Decision Theoretic Analysis," *American Political Science Review*, 68, 525–536.

Fujiwara, T. (2011): "A Regression Discontinuity Test of Strategic Voting and Duverger's Law," *Quarterly Journal of Political Science*, 6, 197–233.

Hamlin, A. and C. Jennings (2011): "Expressive Political Behaviour: Foundations, Scope and Implications," *British Journal of Political Science*, 41, 645–670.

Hansen, S., T. Palfrey, and H. Rosenthal (1987): "The Downsian Model of Electoral Participation: Formal Theory and Empirical Analysis of the Constituency Size Effect," *Public Choice*, 52, 15–33.

Harsanyi, J. (1977): "Morality and the Theory of Rational Behavior," *Social Research*, 44, 623–656.

Herrera, H. and C. Martinelli (2006): "Group Formation and Voter Participation," *Theoretical Economics*, 1, 461–487.

Kawai, K., Y. Toyama, and Y. Watanabe (2021): "Voter Turnout and Preference Aggregation," *American Economic Journal: Microeconomics*, 13, 548–586.

Kawai, K. and Y. Watanabe (2013): "Inferring Strategic Voting," *American Economic Review*, 103, 624–662.

Ledyard, J. O. (1981): "The Paradox of Voting and Candidate Competition: A General Equilibrium Analysis," in *Essays in Contemporary Fields of Economics*, ed. by G. Horwich and J. P. Quirk, Lafayette: Purdue Research Foundation, 54–80.

——— (1984): "The Pure Theory of Large Two-candidate Elections," *Public Choice*, 44, 7–41.

Levine, D. K. and C. Martinelli (2024): "Razor-Thin Mass Elections with High Turnout," *International Economic Review*, 65, 1607–1624.

Levine, D. K. and A. Mattozzi (2020): "Voter Turnout with Peer Punishment," *American Economic Review*, 110, 3298–3314.

Levine, D. K. and T. Palfrey (2007): "The Paradox of Voter Participation? A Laboratory Study," *American Political Science Review*, 101, 143–158.

Lovie, P. (2005): "Coefficient of Variation," in *Encyclopedia of Statistics in Behavioral Science*, ed. by B. S. Everitt and D. C. Howell, John Wiley & Sons.

Morton, R. (1987): "A Group Majority Voting Model of Public Good Provision," *Social Choice and Welfare*, 4, 117–131.

——— (1991): "Groups in Rational Turnout Models," *American Journal of Political Science*, 35, 758–776.

Myerson, R. (1998): "Population uncertainty and Poisson games," *International Journal of Game Theory*, 27, 375–392.

Myerson, R. and R. Weber (1993): "A theory of voting equilibria," *American Political Science Review*, 87, 102–114.

Myerson, R. B. (2000): "Large Poisson Games," *Journal of Economic Theory*, 94, 7–45.

Palfrey, T. and H. Rosenthal (1983): "A Strategic Calculus of Voting," *Public Choice*, 41, 7–53.

——— (1985): "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 79, 62–78.

Palfrey, T. R. and K. Pogorelskiy (2019): "Communication Among Voters Benefits the Majority Party," *Economic Journal*, 129, 961–990.

Pogorelskiy, K. (2024): "Correlated Equilibria in Voter Turnout Games," *Working Paper: University of Warwick*.

Riker, W. and P. Ordeshook (1968): "A Theory of the Calculus of Voting," *American Political Science Review*, 62, 25–42.

Shachar, R. and B. Nalebuff (1999): "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review*, 89, 527–547.

Tullock, G. (1967): *Towards a Mathematics of Politics*, University of Michigan Press.

Uhlaner, C. (1991): "Rational Turnout: The Neglected Role of Groups," *American Journal of Political Science*, 35, 390–422.

III Appendix C: Online Appendix

III.1 Estimation and Implementation Details (Online Appendix)

This appendix explains how we estimate the distribution of election outcomes and how we assess the evidence for group voting.

III.1.1 Sample, parameterization, and likelihood

For election i, let N_i denote eligible population and let $v_{s,i} = N_i\widetilde{\theta}_{s,i}$ be the observed integer vote count on side $s \in \{R, L\}$. The full estimation sample contains 1,724 elections. All models in a comparison use the same election records. Eligible population and the covariates are described in Section 4.1.

We parameterize the cost/value ratios as

$$\log \gamma_{R,i} = X_i^\top \xi_R, \qquad \log \gamma_{L,i} = X_i^\top \xi_L.$$

The group-share index $\log(m_i/N_i) = X_{m,i}^\top \xi_m$ applies before bounding and rounding. The integer group size in the likelihood is

$$m_i = \min\Big\{\lfloor N_i \rfloor, \max\Big\{1, \text{round}\Big[N_i \exp(X_{m,i}^\top \xi_m)\Big]\Big\}\Big\}.$$

The exposure N_i/m_i need not be an integer and does not require an observed partition of the electorate. The equilibrium in Section 4.2 determines the strategic group-arrival intensities $\lambda_{s,i} = N_i\theta_{s,i}/m_i$. We impose $0 < \gamma_{s,i} < 1/2$ and $0 < m_i(\lambda_{R,i} + \lambda_{L,i}) < N_i$. The fitted strategic shares $\theta_{s,i}$ are distinct from the observed shares $\widetilde{\theta}_{s,i}$.

The observation law is

$$\begin{aligned} T_{s,i} &= m_i Y_{s,i} + \varepsilon_{s,i}, \qquad & Y_{s,i} &\sim \text{Poisson}(\lambda_{s,i}), \\ \varepsilon_{s,i} &= P_{s,i}^+ - P_{s,i}^-, \qquad & P_{s,i}^+, P_{s,i}^- &\sim \text{Poisson}(0.005 N_i). \end{aligned}$$

All components and both sides are independent conditional on the covariates. More generally, write $\mu_i = \xi_N N_i$ for the mean of *each* Poisson noise component; the main specification fixes $\xi_N = 0.005$. The difference is symmetric Skellam noise, with mean zero and variance $2\mu_i$. We use the full electorate as strategic exposure; no population is deducted for noise. This additive count law is evaluated on its signed integer support without clipping negative counts or conditioning on nonnegative counts or turnout below N_i.

Let $p(g; \lambda)$ denote the Poisson mass and $q(k; \mu) = \Pr(P^+ - P^- = k)$ the symmetric Skellam mass. Then

$$P_{s,i}(v) = \sum_{g=0}^{\infty} p(g; \lambda_{s,i})\, q(v - m_i g; \mu_i), \qquad v \in \mathbb{Z}.$$

The sum is infinite because negative noise permits $m_i g > v$. The log likelihood is $\sum_i [\log P_{R,i}(v_{R,i}) + \log P_{L,i}(v_{L,i})]$. Its first two moments are

$$\mathbb{E}[T_{s,i}/N_i] = \frac{m_i \lambda_{s,i}}{N_i} = \theta_{s,i}, \qquad \mathrm{Var}(T_{s,i}/N_i) = \frac{m_i^2 \lambda_{s,i} + 2\mu_i}{N_i^2}.$$

Holding the strategic mean share fixed, larger groups increase dispersion. Changing the noise component mean also changes dispersion, but contributes zero to expected votes. At $m_i = 1$, the count law reduces exactly to Skellam$(\lambda_{s,i} + \mu_i, \mu_i)$, giving the independent-voter restriction under the same noise assumption.

III.1.2 Numerical evaluation and optimization

We evaluate the count probabilities using positive series centered on their largest terms, with calculations in log space or scaled relative to the largest term. Both the Skellam mass and its convolution with group arrivals have infinite support. Log-concavity supplies geometric bounds on omitted tails; the default relative truncation tolerance is 2×10^{-14}, apart from floating-point rounding. An independent 128-bit implementation checks the probabilities and likelihoods. We do not substitute a normal approximation, impose a probability floor, or renormalize truncated support.

Integer rounding makes the likelihood flat over regions of the group-index coefficients. We therefore alternate between smooth updates to the log-γ coefficients and searches over the group-size index. Conditional on the log-γ coefficients and fixed noise rule, we tabulate each election's likelihood over candidate integer group sizes. Differential evolution and searches across rounding thresholds propose group-index updates. The implied integer sizes and likelihood are then recomputed from the original coefficients before an update is retained. Conditional on integer group sizes, scaled BFGS updates the log-γ coefficients, using an interior barrier when needed to maintain the strict parameter restrictions. Multiple starting values are compared. The same-sample, same-noise $m = 1$ fit is retained as an attainable likelihood floor for the grouped fit. The reported solution is the best accepted fit found; these searches do not establish a unique global maximum.

III.1.3 Fixed noise and sensitivity checks

The main specification uses $\varepsilon_{s,i} = P^+_{s,i} - P^-_{s,i}$, with independent Poisson components of mean $0.005 N_i$. Table 13 varies this fixed mean and compares Skellam noise with a shifted Poisson. The grouped and $m = 1$ models share the same noise rule within each row.

For shifted-Poisson noise, let

$$f_i = \mu_i - \lfloor \mu_i \rfloor, \qquad P_{s,i} \sim \mathrm{Poisson}(\mu_i), \qquad B_{s,i} \sim \mathrm{Bernoulli}(f_i),$$

and define

$$\varepsilon_{s,i} = P_{s,i} - \lfloor \mu_i \rfloor - B_{s,i}.$$

The Poisson and Bernoulli variables are independent, with independent noise across sides and elections. This integer-valued noise has mean zero and variance $\mu_i + f_i(1 - f_i)$. Skellam noise with component mean μ_i has variance $2\mu_i$. Doubling the shifted-Poisson mean thus approximately matches the Skellam variance, with an additional centering variance of at most $1/4$ per side.

The two models are indistinguishable in the main results.

III.1.4 Bootstrap inference

We use a county-block bootstrap for coefficient uncertainty. A draw samples counties with replacement and includes all observed elections from each sampled county. If a county is drawn twice, its complete election block appears twice. The procedure preserves dependence across elections arising from persistent local organizations, common demographic measurement, or repeated referenda in nearby jurisdictions. The bootstrap requires independence across counties, while allowing dependence among elections within a county.

The main coefficient results use 200 paired county-block bootstrap draws. Within each draw, we re-estimate the grouped and $m = 1$ models at the fixed main noise rule, rebuilding covariate standardization within the resampled full sample and its electorate-size strata. No noise parameter is reselected.

Each appendix sample or covariate specification also uses 200 paired county-block draws, holding its original-sample covariate standardization and the main noise rule fixed. The three size-stratum columns in the appendix coefficient tables use this procedure; their full-sample reference column uses the main bootstrap. The corresponding main bootstrap restandardizes within each resampled stratum. Standard errors of mean group size in Table 13 use 200 county-block draws per noise rule. For the shifted-Poisson rows, we re-estimate the grouped model with original-sample covariate standardization held fixed and calculate mean group size over each draw's resampled elections. These standard errors are conditional on the stated noise rule.

Reported bootstrap standard errors are the standard deviations of the re-estimated quantities across draws. Significance markers on coefficient tables use normal approximations based on those standard errors.

III.1.5 Alternative models and a common count comparison

The alternative models embody different explanations of participation. To compare how well they describe the observed election outcomes, we evaluate probabilities of the same pair of vote counts.

A density of continuous vote shares cannot be compared directly with a probability mass for integer votes: the two have different units. The grouped and $m = 1$ models already provide count probabilities. For an alternative with per-eligible-voter probabilities $(q_{R,i}, q_{L,i})$, the count probability is trinomial over votes against (R), votes for (L), and abstentions:

$$\begin{aligned}\mathrm{Tri}_i(q_R, q_L) &= \frac{N_i^{\mathrm{int}}!}{(N_i\widetilde{\theta}_{R,i})!(N_i\widetilde{\theta}_{L,i})!(N_i^{\mathrm{int}} - N_i\widetilde{\theta}_{R,i} - N_i\widetilde{\theta}_{L,i})!} \\ &\quad \times q_R^{N_i\widetilde{\theta}_{R,i}} q_L^{N_i\widetilde{\theta}_{L,i}} (1 - q_R - q_L)^{N_i^{\mathrm{int}} - N_i\widetilde{\theta}_{R,i} - N_i\widetilde{\theta}_{L,i}},\end{aligned}$$

where N_i^{int} is eligible population rounded to an integer, while $N_i\widetilde{\theta}_{R,i}$ and $N_i\widetilde{\theta}_{L,i}$ are the observed integer vote counts. Every alternative is evaluated on the same observed election records. Differences between the additive count laws and trinomial distributions, including their support restrictions, remain differences between the models rather than differences in the unit of scoring.

Pivotal and expressive voting The pivotal-voter benchmark is based on Coate et al. (2008). The share supporting legalization, μ_i, has a logistic index in Baptist share and metropolitan status. Citizens draw uniform private voting costs; side-specific exponential benefit indices depend on proposition type and jurisdiction characteristics, and a common exponential cost scale depends on weekend elections. We solve the two participation-cutoff conditions at each coefficient vector and maximize the resulting trinomial likelihood. The full-sample adaptation uses expanded proposition indicators, including restaurant-only mixed beverages, and Poisson pivotal probabilities computed from the Skellam difference of the other citizens' votes. Estimation combines Nelder–Mead searches with quasi-Newton refinement, and fitted equilibrium residuals are checked. On the CCM replication sample, Tables 1 and 2 report our re-estimation using the original covariates and approximate pivotal probabilities, followed by evaluation of predictions at those coefficients using finite-population pivotal probabilities.

Our expressive-voting adaptation uses a share s_i supporting legalization and participation propensities $\rho_{R,i}$ and $\rho_{L,i}$, giving $q_{R,i} = (1 - s_i)\rho_{R,i}$ and $q_{L,i} = s_i\rho_{L,i}$. Support and each side's participation have separate logistic indices, including separate weekend coefficients in the participation equations. This differs from the original model's common-cost specification. The support equation includes Baptist share and metropolitan status; participation depends on proposition type, jurisdiction characteristics, and weekend elections. These participation indices contain neither electorate size nor a pivotal probability. Thus the model allows observed characteristics to predict participation without introducing the strategic response to election size. Its coefficients maximize the trinomial likelihood, using a simplex search followed by a quasi-Newton refinement.

Group rule-utilitarian voting The Coate and Conlin (2004) model describes a continuum of voters. It has a Beta-distributed support share and deterministic side-specific turnout conditional on the support realization and the model's common benefit/cost shock. We retain the native share-density estimator and its index structure: the Beta parameters depend on the specified demographic and jurisdiction covariates, benefits depend on proposition type and jurisdiction, and the voting-cost scale depends on weekend elections. We apply a continuity adjustment to the observed vote shares, permitting zero-count elections to enter the native density.

For the count comparison, we fix the common log benefit/cost shock at zero, its median under the implemented standard-normal distribution, and retain the fitted Beta heterogeneity in the share supporting legalization. The resulting benchmark is a median-shock finite-count adaptation. The two turnout rates $(\hat{\sigma}_{R,i}, \hat{\sigma}_{L,i})$ are chosen before support is realized. Consequently, the count probability averages over the fitted support distribution:

$$L_i^{(\mathrm{GRU})} = \int_0^1 \mathrm{Tri}_i((1-\mu)\hat{\sigma}_{R,i}, \mu\hat{\sigma}_{L,i})\, d\hat{F}_i(\mu).$$

This is a continuous mixture, not a trinomial evaluated at mean support. In large elections the integrand can be concentrated in a narrow region, so a small, equally weighted set of support quantiles can miss much of its probability. We use integration centered on the integrand's mode, with bounds on omitted tails, and check the count probabilities against an independent positive finite-sum evaluation. This changes the accuracy of the count evaluation without changing the fitted native-model coefficients. The retained support heterogeneity also distinguishes this benchmark from the expressive and pivotal alternatives: its simulated support share is redrawn for each election replication.

III.1.6 Interpreting the model-comparison tests

The non-nested comparisons ask whether the grouped model has a higher expected log score than a competing model. Let $d_i = \log L_i^{(G)} - \log L_i^{(M)}$ and $\bar{d} = E^{-1}\sum_i d_i$. The unclustered statistic in Table 8 divides $\sqrt{E}\bar{d}$ by the estimated standard deviation of the election-level influence. For expressive and pivotal voting, this influence is the centered score difference under the usual optimized-score conditions. The resulting normal-reference test is an asymptotic comparison under independent elections. It is not a claim that all elections in the data are actually independent.

The GRU comparison requires an additional adjustment because its parameters maximize the native share density rather than the count criterion being compared. Let s_i be the gradient of its native log density, A the negative average native Hessian, and a the average gradient of its

evaluated count log probability. The centered influence used for this comparison is

$$u_i = d_i - \bar{d} - (s_i - \bar{s})^\top A^{-1} a.$$

The last term captures how estimation of the native-model parameters changes the evaluated count score. Omitting it would treat the fitted GRU count probabilities as if they had been estimated by maximizing the count criterion, and would give the wrong variance in general. The point comparison remains the same $\bar{d}$; the correction changes its estimated uncertainty. The generalized normal approximation also requires the native estimator's linearization and regularity of the optimized grouped score.

The county-score bootstrap uses 99,999 resamples of the 216 county clusters, drawing 216 counties with replacement and equal probability in each resample. We retain each selected county's complete election block and hold fitted scores fixed, subtracting the GRU estimation correction above before resampling; this correction is zero for the other alternatives. Let D_c be the corrected score total and E_c the election count in county c. The mean is election-weighted, and its standard error is

$$\widehat{\mathrm{se}}(\bar{d}) = \frac{1}{E}\left[\frac{216}{215}\sum_{c=1}^{216}(D_c - E_c\bar{d})^2\right]^{1/2}.$$

The observed statistic is $\bar{d}/\widehat{\mathrm{se}}(\bar{d})$. In each resample, we recompute the mean and this standard error using the resampled election count. The bootstrap statistic is the resampled mean minus $\bar{d}$, divided by its resampled standard error. The two-sided p-value is one plus the number of bootstrap statistics at least as large in absolute value as the observed statistic, divided by 100,000. Models are not refitted. This asymptotic reference requires independent county clusters and the score-influence regularity conditions.

The $m = 1$ comparison restricts the grouped model, so an ordinary non-nested Vuong reference distribution is inappropriate. Any standardized score difference shown for this restriction is diagnostic only. Integer rounding and nuisance parameters associated with grouping also prevent an automatic standard chi-square calibration of the likelihood-ratio statistic. With the same fixed main noise rule under both hypotheses,

$$T = 2(\ell_G - \ell_1) = 916{,}474.26.$$

We assess this statistic using 99,999 datasets simulated from the fitted $m = 1$ model, with observed covariates and $\xi_N = 0.005$ held fixed. For each side the generating law is Skellam$(\lambda_{s,i} + 0.005N_i, 0.005N_i)$. We retain all simulated signed counts without censoring or refiltering. Writing $\mathbf{t}^*$ for the simulated signed vote counts, we compute a conservative upper bound rather than

optimizing both models in every draw. If η_0 denotes the generating null coefficients, a maximized null likelihood is at least its value at η_0, while a log probability mass cannot exceed zero. Thus

$$T^* \leq -2 \log L_1(\mathbf{t}^*; \eta_0, \xi_N = 0.005).$$

The largest simulated upper bound is 30,603.19, below the observed statistic. There are therefore no possible exceedances among these draws. The add-one Monte Carlo upper-tail estimate is at most 1/100,000 = 0.00001; the one-sided 95% simulation-uncertainty upper limit for the bound's tail probability is approximately 0.000030. These calculations support displaying $p < 0.001$ for this fitted-null reference. They use upper bounds, not re-estimated simulated LR statistics, and do not establish exact size for the composite null or allow arbitrary within-county dependence. Coefficient uncertainty uses the separate county-block bootstrap described above. Throughout the tables, test p-values below 0.001 are displayed as < 0.001 and other p-values are rounded to three decimal places.

Simulated outcomes and additional specifications Simulated turnout and margins show which features of the outcome distribution a model reproduces. For each election, we draw strategic group arrivals and both independent Poisson noise components on each side, forming $T_{s,i} = m_i Y_{s,i} + P^+_{s,i} - P^-_{s,i}$. The main predictive comparisons use 1,000 draws per election, fixed seeds, and equal election weights. Negative counts and turnout above the eligible electorate are retained under the model's unconditioned support. The $m = 1$ model uses the same construction with unit group size. Expressive and pivotal alternatives draw trinomial outcomes. For GRU, we first redraw support from its fitted Beta distribution at the zero log benefit/cost shock and then draw counts, retaining support heterogeneity in both the count comparison and simulations. These draws describe fitted distributions; they do not add independent information to the estimation sample.

Table 16 reports a descriptive two-stage referendum adaptation of Shachar and Nalebuff (1999). First, we regress the log odds of the observed FOR vote share, clipped to $[0.01, 0.99]$, on an intercept, Baptist share, and metropolitan status by ordinary least squares. Applying the logistic function gives fitted support $\hat{s}_i$, and predicted closeness is $1 - |2\hat{s}_i - 1|$. Each side's participation propensity is the exponential of a side-specific covariate index plus common coefficients on predicted closeness and centered $\log N_i$, with this total index clipped to $[-30, 0]$. The probabilities are $q_{R,i} = (1-\hat{s}_i)\rho_{R,i}$ and $q_{L,i} = \hat{s}_i \rho_{L,i}$. We maximize the trinomial likelihood holding the first-stage support predictions fixed. This specification does not solve a leader-effort equilibrium and is excluded from the main model rankings.

III.1.7 Tables

Table 16: Follow-the-Leader Model Adapted to Referenda

Parameter	Estimate
Participation (for): Intercept	−0.5568
Participation (for): City	0.4091
Participation (for): Weekend	−0.3621
Participation (for): All-alcohol/non-mixed	0.0733
Participation (for): Mixed-beverage unrestricted	0.1255
Participation (for): Restaurant-only mixed-beverage	0.0258
Participation (for): new-policy proxy	0.3386
Participation (against): Intercept	−0.4750
Participation (against): City	0.1928
Participation (against): Weekend	−0.2715
Participation (against): All-alcohol/non-mixed	0.0234
Participation (against): Mixed-beverage unrestricted	0.0984
Participation (against): Restaurant-only mixed-beverage	−0.4888
Participation (against): new-policy proxy	0.6164
Effort: predicted closeness	−1.0896
Effort: log(N) (centered)	−0.1041
Support share (first stage): Intercept	0.0810
Support share (first stage): Baptist share	−0.4703
Support share (first stage): MSA	0.1226
Elections	1,724
Count log likelihood	−698,809.4
Simulated mean turnout	0.270
Simulated SD turnout	0.099
Simulated SD margin	0.044
Simulated Pr($\lvert M\rvert > 0.10$)	0.022

Notes: The support-share first stage uses vote-share log odds; remaining coefficients maximize the trinomial count likelihood. Model moments use 1,000 simulated draws per election, with $M = \tilde{\theta}_R - \tilde{\theta}_L$. This benchmark does not use Skellam noise.

III.2 Omitted Proofs (Online appendix)

III.2.1 Omitted details in the proof of Proposition 4

Fix $\vartheta_0 > 0$ and consider $\vartheta = v_I - v_L \leq \vartheta_0$, recalling that $v_I > 0$. In every putative equilibrium in which independent voters mix, $\mathcal{P}_I(\sigma_N, \beta) = c/v_I \geq c/(v_L + \vartheta_0) > 0$. We prove the two uniform bounds used in the main proof.

Step 1. We first show that this equality implies uniform upper bounds on expected turnout. Let N_I be Poisson with mean $T_I^N = \sigma_N \alpha_I N$, and let N_R and N_L be independent Poisson random variables with means $T_R^N = \frac{\alpha_R \beta_R^N N}{m}$, $T_L^N = \frac{\alpha_L \beta_L^N N}{m}$. Conditional on the realization of the group-vote difference, an independent voter is pivotal for at most two adjacent realizations of N_I, each with weight one half. It follows that

$$\mathcal{P}_I(\sigma_N, \beta) \leq \sup_{k \geq 0} \Pr\left\{\text{Pois}(T_I^N) = k\right\}.$$

The largest atom of a Poisson distribution converges to zero as its mean diverges. Therefore, T_I^N is uniformly bounded.

Similarly, using equation (28), $\mathcal{P}_I(\sigma_N, \beta) \leq \sup_{i \in \mathbb{Z}} S_i(T_R^N, T_L^N)$. Moreover,

$$\sup_{i \in \mathbb{Z}} S_i(T_R^N, T_L^N) \leq \min\left\{\sup_{k \geq 0} \Pr\{\text{Pois}(T_R^N) = k\}, \sup_{k \geq 0} \Pr\{\text{Pois}(T_L^N) = k\}\right\}.$$

Consequently, the largest Skellam atom converges to zero whenever $T_R^N + T_L^N \to \infty$. Since $\mathcal{P}_I$ is bounded away from zero uniformly over $\vartheta \leq \vartheta_0$ with $v_I > 0$, there exists a finite constant $\overline{M}$ such that, for every N, every such ϑ, and every equilibrium in which independents mix, $T_j^N \leq \overline{M}$ for $j = I, L, R$.

Step 2. We next establish a uniform lower bound on T_R^N for nonnegative ϑ. We claim that there exist $\vartheta_1 > 0$, $\bar{N}_1 < \infty$, and $\underline{M} > 0$ such that, for every $\vartheta \in [0, \vartheta_1]$, every $N \geq \bar{N}_1$, and every equilibrium in which independents mix, $T_R^N \geq \underline{M}$. Suppose otherwise. Then there would exist sequences $\vartheta_k \in [0, 1/k]$, $N_k \geq k$, and equilibria in which independents mix such that $T_R^{N_k} \longrightarrow 0$. By the uniform upper bound, after passing to a subsequence, $T_I^{N_k} \to T_I^\infty$ and $T_L^{N_k} \to T_L^\infty$. Since $\vartheta_k \to 0$, equation (28) gives

$$\frac{c}{v_L} = \frac{1}{2} e^{-(T_I^\infty + T_L^\infty)}.$$

Because $c/v_L < 1/2$, we have $T_I^\infty + T_L^\infty > 0$. Let N_I^∞ and N_L^∞ be independent Poisson variables with means T_I^∞ and T_L^∞, and set $K = mN_L^\infty + N_I^\infty$. With no other voting R-groups in the limit, an independent is pivotal only when $K = 0$, whereas an R-group is pivotal when $0 \leq K \leq m$. Thus

$$\mathcal{P}_I^\infty = \frac{1}{2}\Pr(K = 0), \qquad \mathcal{P}_R^\infty = \frac{1}{2}\Pr(K = 0) + \sum_{j=1}^{m-1} \Pr(K = j) + \frac{1}{2}\Pr(K = m).$$

The inequality $\mathcal{P}_R^\infty > \mathcal{P}_I^\infty$ is strict: if $T_I^\infty > 0$, then $\Pr(K = 1) > 0$, while if $T_I^\infty = 0$, then $T_L^\infty > 0$ and $\Pr(K = m) > 0$. Hence $\mathcal{P}_R^\infty > c/v_L > c/v_R$. By continuity, the same strict inequality holds along the sequence for all sufficiently large k. But $T_R^{N_k} \to 0$ and $N_k \to \infty$ imply $\beta_R^{N_k} < 1$ eventually, so equilibrium requires $\mathcal{P}_R^{N_k} \leq c/v_R$, a contradiction. This proves the claimed uniform lower bound.

■