The Curse of the Silent Majority*

Marco Battaglini † Thomas R. Palfrey ‡ Gabriel Sekeres §

August 20, 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 reject the hypothesis that groups are unimportant in explaining the data and provide new insights into the social organizations that influence 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 oversimplify 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 I.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 affect turnout, electoral outcomes, and overall welfare. Importantly, we aim at empirically testing our theory to evaluate whether groups are a factor in explaining turnout, and the characteristics of these groups.

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 requires independents to value the outcome almost twice as much as either group. 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. This occurs because, as we will explain in greater detail, essentially *all* models of voting with exclusively independent voters (including behavioral models) predict a standard deviation of the margin of victory, σ_{MV}, approximately equal to $\sqrt{\mu}$, the square root of the expected turnout. In elections with high turnout, standard models therefore predict a small standard-deviation-to-mean ratio, $\sigma_{MV}/\mu = 1/\sqrt{\mu}$, and hence a thin-tailed distribution of the margin of victory.[4] This does not happen in our model with organized groups, where the presence of groups prevents tail probabilities from collapsing.

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 generates out-of-sample predictions that are better by orders of magnitude. Because the model with groups nests the model without groups as a special case, we can formally reject the hypothesis that groups can be ignored, even after accounting for the additional parameter. Although neither the net expected benefit of voting nor the size of organized groups is directly observed, we prove that both are fully identified from the data. 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.

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, 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

[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 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.

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 (1987)). 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 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

[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.

[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

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 opposite side 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 depend 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 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

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

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. (2017), 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. (2017) 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 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.

[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.

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 $m > 1$ 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 $l = I, L, R$). 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 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

[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.

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 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 have three natural features: they are easily enforceable, treat all group members symmetrically, and do not require the group to send private messages to individual members. 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 of a type j group vote for their favorite candidate, and $1-\beta_j$ is the probability that all m members abstain. 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.

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:

[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.

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 $P^L(c) \geq 1 - \delta$ in a large election.*

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:

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

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

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

[13] To simplify the discussion here and only for the sake of the illustration, we are writing these pivot probabilities using the approximations developed in Myerson (2000) that converge to the true value in elections with large participation. The actual proof of Proposition 2 does not use approximations and proves the result in full generality. The exact formulas for the pivot probabilities are given in the appendix, equations 34 and 35.

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:

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

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:

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

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_L and v_R, 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 5.5.2, 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, the cost of voting cannot be too close to $v_R/2$, the point at which a voter has a dominant strategy to 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 $\bar{c} = \frac{v_R}{2} e^{1-\frac{v_R}{v_L}}$.[14]

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 $1 > \beta_R > \frac{\alpha_L}{\alpha_R}\beta_L > 0 = \sigma$ and 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:

[14] If $v_R = v_L$ then $\bar{c} = v_R/2$.

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. See Appendix. □

Proposition 3 has two lessons. First, when c is small, Δ^* is pinned down as v_L/v_R. 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 L wins with probability approaching one as $n \to \infty$ in all equilibrium sequences.

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. *In a large election with organized groups, there exists a ϑ^*, such that for $\vartheta < \vartheta^*$, $\sigma = 0$ in all equilibria.*

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 v_I to be sufficiently close to v_L.

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 any model in which voters act independently of each other, regardless of the specific source of motivation. Thus, our argument for incorporating groups in the theory of voting applies equally well to 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

Using the large-election approximations introduced above, the equilibrium conditions can be approximated by the following two equations:[15]

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

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

where N_i and v_i are the size and value of voters who prefer candidate i. To interpret these conditions, 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, Proposition 1 tells us that the voters who support L is the class of voters j with the highest v_j: so we have $v_L = v_I$ and $N_L = \alpha_I n$ when $v_L \leq v_I$; and $v_L = v_L$ and $N_L = \alpha_L n$ when $v_L > v_I$. 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

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

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 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:

$$\frac{e^{-\left(1-\frac{\gamma_L}{\gamma_R}\right)^2 \frac{n}{m}\theta_R}}{4\left(\pi \frac{\gamma_L}{\gamma_R}\frac{n}{m}\theta_R\right)^{\frac{1}{2}}} = \frac{\gamma_L}{\left[1+\frac{\gamma_L}{\gamma_R}\right]} \tag{9}$$

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

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 $\widetilde{\mu}_j$

[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. These cases are however impossible when n is large, so ignored here for simplicity.

[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. We note here that under the assumption of Proposition 3—that is, $v_I \leq v_R + v_L$ for large n—independents abstain with probability one. This implies that we need $\alpha_R n \beta_R \geq \widetilde{\mu}_R$. Since $\beta_R \leq 1$, it follows that we have a well defined lower bound for α: $\alpha_R \geq \widetilde{\mu}_R / n$.

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 predicted expected turnout of groups for candidate j is a Poisson with mean and variance:

$$\mu_j = m \left\lceil \frac{\alpha_j n}{m} \right\rceil \beta_j \simeq n\theta_j \tag{11}$$

$$\sigma_j^2 = m^2 \left\lceil \frac{\alpha_j n}{m} \right\rceil \beta_j \simeq mn\theta_j = m\mu_j \tag{12}$$

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 over θ_L and ignoring the integer approximation, from (10) we obtain:

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

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:

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

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

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

[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.

Table 1: CCM-Sample Turnout by Electorate Size

Panel	N	Data	CCM	$m = 1$	Grouped
$n < 218$	35	0.650	0.699	0.668	0.641
$218 \leq n < 353$	35	0.556	0.535	0.553	0.546
$353 \leq n < 520$	35	0.509	0.500	0.509	0.512
$n \geq 520$	36	0.434	0.428	0.441	0.450

Notes: Panels are empirical quartiles of the CCM-sample eligible-voter distribution with cutoffs 218.1, 353.4, and 520.2. The CCM column re-estimates the finite pivotal-voter model with the Coate et al. (2008) covariates. Turnout is $(V_R + V_L)/n$.

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 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 (i.e., $n < 218$), medium (i.e., $n \in [218, 353)$ and $n \in [353, 520)$), and large (i.e., $n \geq 520$) electoral district, along with the corresponding predictions by the pivotal voter model estimated by CCM and the model estimated by us that will be discussed in detail in Section 4.[21] As can be seen the models tend to estimate these values reasonably well. For instance, for the case of cities with $n \in (520, 900)$ voters, expected turnout is about 43%, CCM predicts it to be about 43% and our model with $m = 1$ predicts it to be about 44%.[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

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

[21] This breakdown of the sizes of electoral districts is taken from Coate et al. (2008), who restrict the analysis to districts with fewer than 900 citizens. In Section 4, we will show that their analysis is robust to extending the dataset over time and to districts of all sizes.

[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 have been recomputed using the same dataset code they use in their analysis. A detailed discussion of the

Table 2: CCM-Sample Tail Probabilities by Electorate Size

	Data	CCM	$m = 1$	Grouped
$n < 218$, $N = 35$				
$\Pr(\lvert\text{margin}\rvert > 0.05)$	0.800	0.523	0.586	0.831
$\Pr(\lvert\text{margin}\rvert > 0.10)$	0.629	0.215	0.294	0.681
$\Pr(\lvert\text{margin}\rvert > 0.15)$	0.514	0.077	0.139	0.550
$\Pr(\lvert\text{margin}\rvert > 0.20)$	0.486	0.029	0.064	0.426
$\Pr(\lvert\text{margin}\rvert > 0.25)$	0.343	0.012	0.030	0.322
$218 \leq n < 353$, $N = 35$				
$\Pr(\lvert\text{margin}\rvert > 0.05)$	0.886	0.332	0.428	0.737
$\Pr(\lvert\text{margin}\rvert > 0.10)$	0.829	0.051	0.093	0.512
$\Pr(\lvert\text{margin}\rvert > 0.15)$	0.600	0.005	0.011	0.329
$\Pr(\lvert\text{margin}\rvert > 0.20)$	0.429	0.000	0.001	0.196
$\Pr(\lvert\text{margin}\rvert > 0.25)$	0.343	0.000	0.000	0.108
$353 \leq n < 520$, $N = 35$				
$\Pr(\lvert\text{margin}\rvert > 0.05)$	0.771	0.216	0.471	0.709
$\Pr(\lvert\text{margin}\rvert > 0.10)$	0.457	0.010	0.131	0.458
$\Pr(\lvert\text{margin}\rvert > 0.15)$	0.257	0.000	0.013	0.261
$\Pr(\lvert\text{margin}\rvert > 0.20)$	0.086	0.000	0.001	0.141
$\Pr(\lvert\text{margin}\rvert > 0.25)$	0.057	0.000	0.000	0.067
$n \geq 520$, $N = 36$				
$\Pr(\lvert\text{margin}\rvert > 0.05)$	0.722	0.097	0.790	0.764
$\Pr(\lvert\text{margin}\rvert > 0.10)$	0.417	0.001	0.338	0.520
$\Pr(\lvert\text{margin}\rvert > 0.15)$	0.222	0.000	0.051	0.313
$\Pr(\lvert\text{margin}\rvert > 0.20)$	0.139	0.000	0.003	0.161
$\Pr(\lvert\text{margin}\rvert > 0.25)$	0.083	0.000	0.000	0.070

Notes: Panels are empirical quartiles of the CCM-sample eligible-voter distribution with cutoffs 218.1, 353.4, and 520.2. The CCM column re-estimates the finite pivotal-voter model with the CCM covariates. Model entries use 250 simulations per election.

is very significant. For instance, margins of victory above 20% are common in the data, but essentially never occur in the estimated models. For districts with more than 520 citizens, the probability of a margin of victory larger than 20% is still nearly 14%; it is essentially zero in both the CCM model and our model. 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:

Texas Liquor referenda is presented in Section 4.

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. These variables can be uniquely found solving (13) and (15), and give us a model that by construction perfectly estimates turnout since this statistic is used to calibrate γ_R and γ_L. The model predicts that support for option j is a Poisson with mean μ_j. Turnout is therefore predicted to be distributed according to a Poisson with mean $\mu = \mu_R + \mu_L$ and standard deviation $\sigma_T = \sqrt{\mu_R + \mu_L}$. The margin of victory is the difference in turnout between R and L, therefore it is predicted to be a random variable distributed according to a Skellam $S_i(\mu_R, \mu_L)$ with mean $\mu_{MV} = \mu_R - \mu_L$ and standard deviation $\sigma_{MV} = \sqrt{\mu_R + \mu_L}$.[24]

Using these observations, 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

[24] In general, given two 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 $\mu_1 + \mu_2$.

Table 3: National Two-Candidate Elections, 1980–2024: Eligible-Voter Margin Tails

Statistic	All offices	US House	US Senate
N	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.

different model adopted by Coate et al. (2008) share the same property). If instead of assuming a Poisson population of voters as in Myerson (2000), one instead assumed a classic Palfrey and Rosenthal (1983) model with a fixed constant n of voters, very similar results obtain. In that case, for large n the distribution of turnout for candidate j would be approximately a Binomial with mean μ_j, and variance

$$\frac{\mu_j}{n}(1 - \frac{\mu_j}{n})n \simeq \mu_j,$$

for large n, just as in the Poisson case (since both the Binomial and the Poisson converge to a Normal as $n \to \infty$). The difference $\mu_R - \mu_L$ would be approximately a normal with mean $\mu_R - \mu_L$, and variance

$$\frac{\mu_R}{n}(1 - \frac{\mu_R}{n})n + \frac{\mu_L}{n}(1 - \frac{\mu_L}{n})n \simeq \mu_R + \mu_L,$$

once again just as in the Poisson case (since the Skellam also converges to a Normal as $n \to \infty$).[25] The results, therefore, would be very similar.[26] In both 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

[25] More generally, the approximation of the variance by the mean requires the individual voting probability to be small, but it is not needed for the result: since $\mathrm{Var}(V_j) = \mu_j(1 - \mu_j/n) \leq \mu_j$, we still have $\sigma_j/\mu_j \leq 1/\sqrt{\mu_j} \to 0$, and, similarly, $\sigma_{MV}/\mu \leq 1/\sqrt{\mu} \to 0$.

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

this, consider a different theory of election that makes the prediction that voters for j 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 turnout, then $\beta_j(\gamma, X_k) = \widetilde{\mu}_j$. Then, as an implication of independent choice behavior, we again obtain the prediction that the Margin of Victory is a the difference of two Poissons (or Binomials) with mean $\mu_{MV} = \widetilde{\mu}_R - \widetilde{\mu}_L$ and standard deviation $\sigma_{MV} = \sqrt{\widetilde{\mu}_R + \widetilde{\mu}_L}$, thus with the property that $\sigma_{MV}/\mu = 1/\sqrt{\mu} \simeq 0$ for large μ.[27]

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,

[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.

[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.

Table 4: National Election Moment Scaling

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 equal-count 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 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.

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 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 by 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 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

[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.

[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.

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. 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 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 $\left\lceil \frac{\theta_j n}{m_n} \right\rceil$, where m_n is the size of the groups in district of size n. Turnout for candidate j is $\mu_j(n) = m_n \left\lceil \frac{\theta_j n}{m_n} \right\rceil$, since each group votes with probability β_j, and each group has m_n members. Similarly to the classical case of $m_n = 1$, total expected turnout is:

$$\mu_n = \mu_{n,L} + \mu_{n,R} = \left\lceil \frac{\theta_L n}{m_n} \right\rceil m_n + \left\lceil \frac{\theta_R n}{m_n} \right\rceil m_n \simeq \theta_L n + \theta_R n \tag{16}$$

which is essentially independent of m_n, except for integer rounding. The variance of turnout, however, is:

$$\sigma_n^2 = Var(m_n N_R + m_n N_L) = m_n^2 Var(N_R + N_L) = m_n^2 \left[\left\lceil \frac{\theta_L n}{m_n} \right\rceil + \left\lceil \frac{\theta_R n}{m_n} \right\rceil \right] \simeq m_n \left[\mu_{nR} + \mu_{n,L} \right] \tag{17}$$

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

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

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 $\sqrt{m_n/\left(\mu_{n,R}+\mu_{n,L}\right)}$: 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. We only note here that a desirable property of these referenda is that they are mostly held separately from other elections. Texas election law mandates that major elections (presidential, gubernatorial, and congressional elections) are held on prespecified "uniform election dates". Liquor referenda are instead held outside these dates. Coate and Conlin (2004) observe that occasionally other issues are voted on the same dates, but this occurs sporadically. This guarantees that liquor referenda are essentially single-issue elections in which turnout is not affected by the presence of other races.

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" is always liberalizing and a vote "against" is always

maintaining the *status quo*.[31] Considering data from 1937 to 2019, this gives us a total of 1,703 referenda, comprising districts from 102 to 770,691 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 this variable from external demographic sources. For county and justice precinct 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 1971 onward, and the population aged 25 and above plus four-fifths of the population aged 20–24 through 1970 (the voting age was 21 until the 26th amendment's ratification in 1971). 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.[33]

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.[34] Our baseline analysis uses

[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.

[33] 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.

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

Table 5: Summary Statistics: Main Structural Estimation Sample

	Full ($N = 1{,}703$)	Small ($N = 575$)	Medium ($N = 592$)	Large ($N = 536$)
Eligible voters (n)	13,667 (45,836)	779 (506)	4,763 (1,883)	37,327 (76,501)
Votes for	756 (1,815)	135 (102)	436 (414)	1,775 (2,951)
Votes against	785 (1,813)	135 (118)	491 (477)	1,806 (2,932)
Turnout rate	0.247 (0.190)	0.391 (0.182)	0.205 (0.150)	0.140 (0.138)
Margin of victory	−0.005 (0.096)	0.002 (0.139)	−0.009 (0.071)	−0.006 (0.050)
Pct. Baptist	0.276 (0.110)	0.290 (0.107)	0.281 (0.115)	0.255 (0.103)
Avg. SBC congregation size	529 (270)	477 (219)	479 (244)	641 (311)
City election	0.532 (0.499)	0.920 (0.272)	0.400 (0.490)	0.261 (0.440)
Weekend election	0.623 (0.485)	0.638 (0.481)	0.601 (0.490)	0.631 (0.483)
Alcohol (off-prem)	0.356 (0.479)	0.407 (0.492)	0.363 (0.481)	0.293 (0.456)
Alcohol (on/off)	0.201 (0.401)	0.186 (0.390)	0.223 (0.417)	0.192 (0.394)
For-side wins	0.466 (0.499)	0.537 (0.499)	0.427 (0.495)	0.431 (0.496)

Notes: Cells report means with standard deviations in parentheses; binary variables are reported as proportions. Baseline sample: previously dry local-option elections with non-missing key covariates, turnout no greater than one, and at least 100 eligible voters. Eligible voters min/median/max: Full: 102 / 4,335 / 770,691; Small: 102 / 669 / 1,998; Medium: 2,006 / 4,510 / 8,494; Large: 8,503 / 16,721 / 770,691. Small, medium, and large jurisdictions are defined by manual electorate cutoffs into approximate terciles: $n < 2000$, $2000 \leq n < 8500$, and $n \geq 8500$.

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, $n < 2{,}000$; medium, $2{,}000 \leq n < 8{,}500$; and large, $n \geq 8{,}500$. 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, an indicator for weekend elections, and indicators for the two broad alcohol proposition types. The grouped model uses the same turnout and vote-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 robustness checks – restricting to the first referendum held in each district, using alternative electorate-size thresholds, and a richer set of covariates, as well as splitting the elections into strata based on year. All results are qualitatively identical.

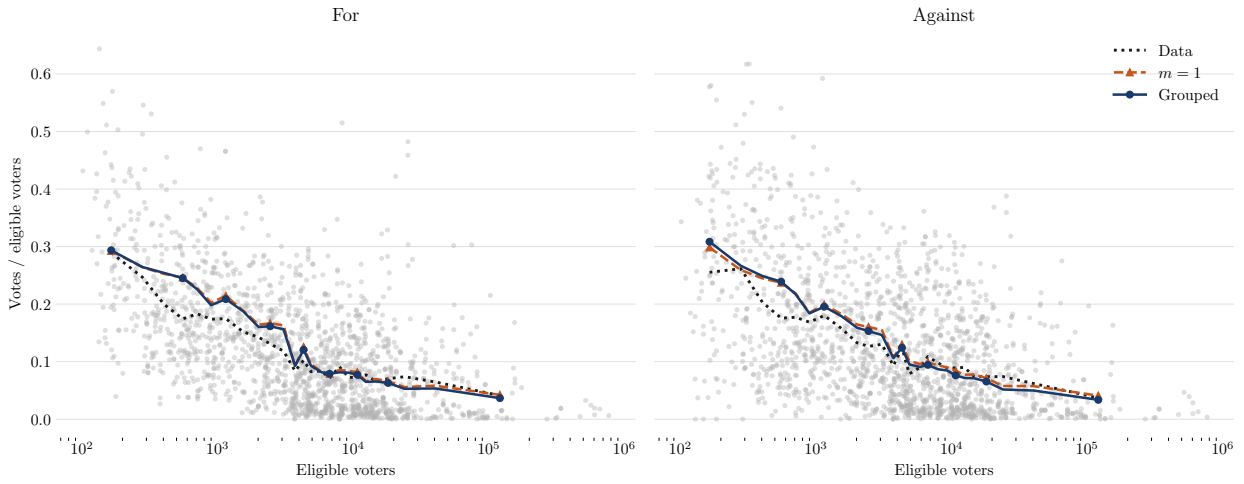


Figure 1: Turnout by electorate size in the Texas main sample. The two panels report votes for (left) and against (right) as a fraction of eligible voters; in each panel the empirical pattern is compared with fitted turnout from the $m = 1$ benchmark and the grouped model.

4.2 Structural Estimation

As discussed in Section 3.2.1, if we had an arbitrarily rich dataset, we would be able to estimate the model non-parametrically, by simply computing, in each electoral district, the mean and variance of the votes for and against legalization over elections held over time in the district. With the available dataset, however, we need a more parsimonious specification, that explicitly models the key unobserved variables $\gamma_{L,i}$, $\gamma_{R,i}$ and m_i for each electoral unit i as a function of observables.

To this goal, it is first convenient to do a simple change of variables. For election i, let n_i be eligible voters, let $V_{R,i}$ and $V_{L,i}$ be the votes for and against legalization, V_i be total turnout $V_i = V_{L,i} + V_{R,i}$, and $t_i = V_i/n_i$ be the turnout rate. Moreover, let X_i denote the baseline covariates described in Section 4. Rather than estimating $(\gamma_{R,i}, \gamma_{L,i})$ directly for each election i, we estimate (i) the ratio r_i pinned down by (13) and (ii) (fractional) total turnout t_i:

$$r_i \equiv \frac{\gamma_{L,i}}{\gamma_{R,i}} = \sqrt{\frac{\theta_{R,i}}{\theta_{L,i}}} \qquad \text{and} \qquad t_i \equiv \theta_{R,i} + \theta_{L,i} \tag{18}$$

for each election i. Given any (r_i, t_i) with r_i and t_i, the implied turnout intensities $(\theta_{R,i}, \theta_{L,i})$ are obtained by solving the system in (18). Writing $r_i^2 = \theta_{R,i}/\theta_{L,i}$ and $t_i = \theta_{R,i} + \theta_{L,i}$ gives

$$\theta_{Ri} = \frac{r_i^2 t_i}{1 + r_i^2} \qquad \theta_{Li} = \frac{t_i}{1 + r_i^2} \tag{19}$$

In the benchmark specification with only individual voters and no groups, $m = 1$, we only need

to model (r_i, t_i). We do it by modeling r_i and t_i as follows:[35]

$$\log r_i = X_i^\top \xi_r \text{ and } t_i = \frac{\exp(X_i^\top \xi_t)}{1 + \exp(X_i^\top \xi_t)}$$

Given this, the two vote totals are independent Poisson random variables, $V_{R,i} \sim \text{Poisson}(n_i \theta_{R,i})$ and $V_{Li} \sim \text{Poisson}(n_i \theta_{L,i})$, so the election-level likelihood contribution is:

$$L_i^{(1)}(\xi_r, \xi_t) = p(V_{R,i}; n_i \theta_{R,i})\, p(V_{L,i}; n_i \theta_{Li}) \tag{20}$$

where $p(v; \lambda)$ is the Poisson probability mass function. This is the direct analogue of the independent-voter benchmark in Coate et al. (2008). We estimate this model by solving for the ξ_r, ξ_t that maximize the sum of the logs of (20) over elections.

For the more general model with organized groups, we also need to model m_i, i.e. the size of groups in each election i. It is convenient to model the size of the groups as a share of the total population:

$$m_i / n_i = e^{X_{m,i}^\top \cdot \xi_m}$$

where $X_{m,i}$ contains an intercept, $\log n_i$, $[\log(n_i)]^2$, city status, Baptist share, and standardized average Southern Baptist congregation size. The exponential index is a continuous latent group-share index. In the likelihood, we restrict group size to the feasible interval $[1, n_i]$ and round to the nearest integer m_i.

To estimate the model with groups we need to confront a limitation of the theoretical model, which to be manageable assumes all groups in an election i have size m_i. This implies that the data generated by the model will be a multiple of m_i, which the observed data typically are not. We reconcile the data with the model by assuming that the discrepancy between the observed turnout and the theoretical prediction is due to individuals either failing to arrive when their group decides to vote, or erroneously voting when their group decides to stay home. Formally, fix some $\kappa \in (0, 1)$ and let $B_i = \lfloor n_i / m_i \rfloor$ denote the number of groups. The theory predicts that the number of groups that will arrive for side s in election i is $B_i \cdot \theta_{s,i}$. Each mobilized group delivers $Z \sim \text{Binomial}(m_i, (1 - \kappa))$ votes rather than m_i exactly, and individual voters[36] arrive randomly

[35] This reparameterization is convenient numerically in the estimation. When $\theta_{R,i}$ and $\theta_{L,i}$ are estimated directly, estimation noise will sometimes send the estimated $\hat{\theta}$ values into ranges where the system of equations (9) and (10) do not admit interior solutions, flattening the objective function. The intuition is that there are potential values of $\theta_{R,i}$ and $\theta_{L,i}$ which cannot be justified by any cost/value pair. For example, if one side entirely turns out and the other stays at home, someone on the side which turns out would rather stay home. These corner regions can be avoided by estimating ratio and turnout directly and then recovering the implied $\theta_{R,i}$ and $\theta_{L,i}$, which will always be (weakly) interior.

[36] Note that this is distinct from the independent voters not organized into groups – we are assuming here that these are people who, per the theory, *should not* vote but do so anyway. Most plausibly, we can think of these voters as recipients of individual-level cost shocks that are unobserved.

at the same rate as the expected shortfall: assume the mass of individual voters arriving for side s in election i is $U_{s,i} \sim \text{Poisson}(\kappa\, m_i B_i \theta_{s,i})$. We have that

$$V_{s,i} = \sum_{j=1}^{Y_{s,i}} Z_{j,s,i} + U_{s,i}, \qquad \mathbb{E}\left[V_{s,i}\right] = m_i B_i \theta_{s,i}.$$

Two properties are nice. First, this is mean-preserving both on the aggregate election level and for each side – the expected number of voters remains the same, because κ is fixed globally. Second, fixing $m = 1$ the two components of V_s collapse into $V_s \sim \text{Poisson}(\theta_s \cdot n)$ for any κ, so the grouped model nests the $m = 1$ model exactly. The estimation of κ is discussed in detail below and in Appendix II.

Conditional on the data $(X_i, X_{m,i})$ and the parameters $\xi = (\xi_r, \xi_t, \xi_m)$, the model with groups assigns the following probability to a vote count v:

$$\Pr(V_{s,i} = v) = \sum_{k \geq 0} p(k; B_i \theta_{s,i}) \sum_{j=0}^{\min\{km_i, v\}} B(j; km_i, 1-\kappa)\, p(v-j; \kappa m_i B_i \theta_{s,i}), \tag{21}$$

where $p(k; x)$ is the Poisson probability of k at mean x and $B(j; N, q)$ is the binomial probability of j successes in N trials with success probability q. Because a mobilized group can deliver fewer than m_i votes, the outer sum is not bounded by $\lfloor v/m_i \rfloor$; we truncate it at a data-dependent bound chosen so the omitted tail is numerically negligible. The election-level likelihood contribution is therefore

$$L_i^{(G)}(\xi_r, \xi_t, \xi_m; \kappa) = \Pr(V_{R,i} = v_{R,i}) \cdot \Pr(V_{L,i} = v_{L,i}), \tag{22}$$

where each probability is evaluated using (21). We evaluate the inner convolution exactly with streaming log-sum-exp recurrences for numerical stability (replacing it with a moment-matched normal kernel only when km_i exceeds a large fixed budget, which we verify leaves the likelihood unchanged to numerical precision), and maximize the sample log likelihood

$$\ell(\Theta) = \sum_i \log L_i(\Theta) \tag{23}$$

by Nelder–Mead.

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 $m \geq 1$, thereby allowing for the presence of organized groups. We emphasize three main sets of results. First, we show that the model with groups can account for both expected turnout and the distribution of the margin of victory. This improvement is not simply due to the addition of an extra parameter. Because the model without groups is nested as a special case of the general model, we can use a standard likelihood-ratio test to formally reject the hypothesis that groups are unimportant for explaining the data. Second, we examine what the estimates reveal about the nature of organized groups and derive bounds on the share of the electorate that must be organized for the model to rationalize the observed outcomes. Third, we provide a broader comparison with alternative, non-nested models of voting that have been estimated using the same data and show that our model performs very well relative to them. Sections 5.3 and 5.4 summarize a battery of robustness checks and out-of-sample validation exercises, which are reported in full in Appendix II.

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 that voting costs are publicly observed and that the number of voters follows a Poisson distribution, as in Myerson (1998). In this section, we show that, despite these differences, the two models generate remarkably similar qualitative and quantitative results. This similarity provides a useful validation of both approaches.

Tables 1 and 2, that were briefly discussed in Section 3.2.2, compare the estimates in CCM with our estimates of turnout and the margin of victory using data from the original CCM paper. We return to them here to highlight that CCM and our model have a very similar performance using the original dataset in CCM: 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, the probability of a margin of victory larger than 15%, has an empirical probability of over 22%: the CCM model, however, predicts essentially a zero probability for this event; and our

Table 6: Turnout and Tail Probabilities by Size Group. The small, medium, and large strata are defined by $n < 2{,}000$, $2{,}000 \leq n < 8{,}500$, and $n \geq 8{,}500$ eligible voters.

Stratum	Series	Turnout	Pr(|margin| > 0.05)	Pr(|margin| > 0.10)	Pr(|margin| > 0.15)	Pr(|margin| > 0.20)
Full	Empirical	0.247	0.402	0.209	0.114	0.062
	CCM	0.297	0.049	0.010	0.002	0.000
	m=1	0.278	0.113	0.010	0.001	0.000
	Grouped	0.268	0.371	0.168	0.087	0.047
Small	Empirical	0.391	0.668	0.398	0.237	0.148
	CCM	0.529	0.146	0.029	0.006	0.001
	m=1	0.387	0.291	0.045	0.006	0.001
	Grouped	0.404	0.629	0.351	0.189	0.101
Medium	Empirical	0.205	0.326	0.142	0.073	0.027
	CCM	0.230	0.000	0.000	0.000	0.000
	m=1	0.206	0.059	0.002	0.000	0.000
	Grouped	0.213	0.310	0.062	0.010	0.002
Large	Empirical	0.140	0.200	0.080	0.028	0.007
	CCM	0.124	0.000	0.000	0.000	0.000
	m=1	0.146	0.035	0.000	0.000	0.000
	Grouped	0.126	0.091	0.005	0.000	0.000

Notes: $N = 1{,}703$ elections (575 small, 592 medium, 536 large). Model moments are computed from simulated elections using the main fitted specifications. The CCM row re-estimates the Coate–Conlin–Moro finite pivotal-voter model on the same main sample by gradient-free iterated Nelder–Mead. Margin of victory is $(V_R - V_L)/n$ in the data and the simulated vote difference scaled by the election's eligible-voter count in the simulations.

model with $m = 1$, predicts a probability equal to 4.7%.

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 that both pivotal-voter models capture the sharp decline in turnout as electorate size increases. Average turnout in the data falls from 39.1% in small jurisdictions to 20.5% in medium jurisdictions and 14.0% in large jurisdictions; both models reproduce this downward gradient: the CCM model predicts turnout rates of 52.9%, 23.0%, and 12.4%, respectively, while our restricted model with $m = 1$ predicts 38.7%, 20.6%, and 14.6%. 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 ($2{,}000 \leq n < 8{,}500$) they occur with probability 32.6% in the data, while the $m = 1$ benchmark predicts 5.9% and CCM predicts essentially zero. In large jurisdictions ($n \geq 8{,}500$), these margins occur with probability 20.0% in the data, while the $m = 1$ benchmark predicts 3.5% and CCM again predicts zero. Even among small elections ($n < 2{,}000$), margins

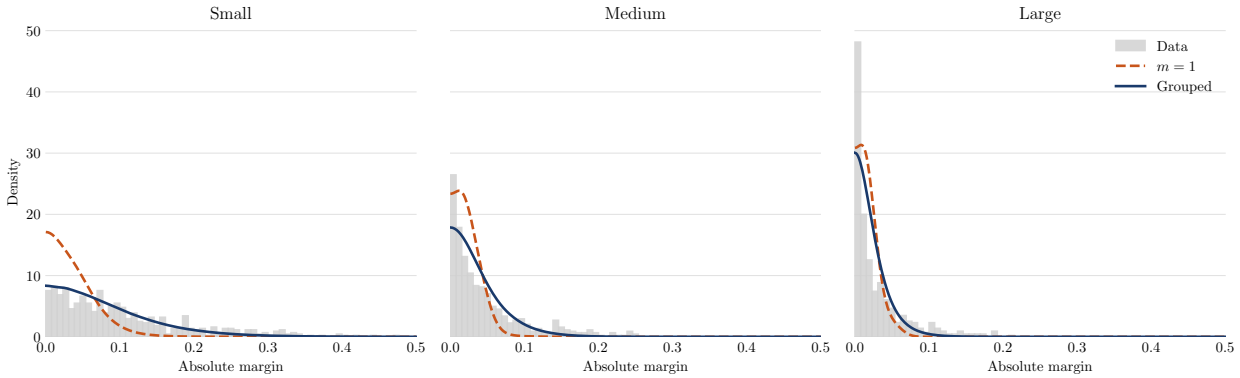


Figure 2: Absolute margin of victory by electorate-size stratum. Panels report, from left to right, small, medium, and large jurisdictions; in each panel the empirical distribution is compared with simulations from the fitted $m = 1$ benchmark and, anticipating the results of the next section, from the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

of victory above 15% occur with probability 23.7% in the data. By contrast, they occur with probability 0.6% in the $m = 1$ benchmark and 0.6% 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).

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 59, with a median of 51.

Comparing the estimates with the case discussed above, a number of lessons emerge clearly. The first lesson is that we can confidently reject the null hypothesis that organized groups do not improve the fit of the pivotal voter model.[37] After allowing for groups, the full-sample log likelihood rises from −1,023,634 under $m = 1$ to −36,814 under the grouped model. Naturally, it is not surprising that the model improves the fit once an additional parameter is introduced. Still, the improvement

[37] We return to the comparison of alternative models in Section 5.2.3, where we provide additional formal tests and further analysis.

Table 7: Main Grouped-Vote Estimates

	Vote ratio $\log r$	Turnout $\text{logit}(t)$	Group share $\log(m/n)$
Intercept	-0.114^{***}	0.779^{***}	1.922^{***}
	(0.017)	(0.047)	(0.016)
$\log(n)$	0.015^{***}	-0.353^{***}	-0.805^{***}
	(0.003)	(0.004)	(0.002)
$[\log(n)]^2$			0.007^{***}
			(0.000)
Baptist share	-0.175^{**}	2.396^{***}	-0.067^{**}
	(0.072)	(0.165)	(0.029)
City	0.174^{***}	0.864^{***}	-0.287^{***}
	(0.023)	(0.075)	(0.008)
Weekend	-0.032	-0.221^{***}	
	(0.023)	(0.022)	
Alcohol (off-prem)	-0.128^{***}	-0.118^{***}	
	(0.026)	(0.038)	
Alcohol (on/off)	-0.049	0.059	
	(0.035)	(0.052)	
SBC congregation size (std.)			0.116^{***}
			(0.016)
N		1,703	
Log likelihood		-36813.8	
Mean $\hat{m}$		59.12	
Median $\hat{m}$		51.00	

Notes: Full-sample maximum-likelihood estimates for the grouped-vote model. The leakage parameter is fixed at $\kappa = 0.069$. A blank cell indicates that the covariate does not enter that equation; the group-share equation uses standardized Southern Baptist average congregation size in addition to the baseline covariates. County-block bootstrap standard errors (200 draws) are in parentheses. *, **, and *** denote significance at the 10, 5, and 1 percent levels. Estimates by electorate-size stratum are reported in Table 9.

in fit more than compensates for the larger parameter space. The likelihood-ratio statistic for the restriction $m = 1$ is 1,973,641: the nominal likelihood-ratio test rejects the no-groups restriction decisively. Because elections within a county are not independent, we do not lean on the nominal χ^2 distribution alone: the same conclusion follows from the bootstrap distribution of the group-size coefficients in the group-share equation is tightly concentrated away from zero.

A second indication of the quality of the estimates is that the coefficients are economically sensible across all three equations, both for the full sample (Table 7) and for the estimates on subsamples of three size categories, which are reported in Appendix Table 9. Moreover, the estimates align closely with the institutional interpretation of these elections as contests organized around religious congregations and local civic institutions. In the vote-ratio equation, support for liberalization decreases with the county's Baptist population share and increases in citywide elections,

while the coefficient on electorate size is essentially zero. Thus, preferences over liquor liberalization do not vary systematically with the size of the jurisdiction, but they do depend on the religious composition of the electorate. In the turnout equation, the estimated elasticity of turnout with respect to electorate size is -0.35, consistent with the central comparative static of the pivotal-voter model. The Baptist population share strongly increases mobilization on both sides of the issue: at the sample mean, a one-standard-deviation increase in the Baptist share, equal to 0.11, raises predicted turnout by approximately seven percentage points. Weekend elections, by contrast, reduce turnout. The group-size equation captures the model's central structural mechanism.

The estimates imply that group size m increases with the electorate, but much less than proportionally. The fitted elasticity of group size with respect to electorate size is approximately 0.3 at the median electorate and rises to about 0.4 in the largest jurisdictions. The model therefore does not require implausibly large groups to fit the data. Groups are larger relative to the electorate in cities and in counties with higher Baptist population shares. Most directly, the average size of Southern Baptist congregations enters positively: where congregations are larger, the estimated units of coordination are also larger. These coefficients are precisely estimated, and the county-block bootstrap standard errors indicate that all group-share loadings are statistically significant at the one-percent level.

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 40.2% in the data: it is only 4.9% in the CCM model and 11.3% in our model with $m = 1$, but 37.1% 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 between 102 and 1,998 eligible voters. It also performs

relatively well for medium districts, with eligible voters between 2,006 and 8,493: in the data, margins of victory larger than 10% occur with probability equal to 14.2%; in the model with $m = 1$, the probability is 0.2%; in the model with estimated m, the probability is 6.2%. 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 8.0% in these elections. The $m = 1$ model predicts essentially zero probability for these events, whereas the grouped model predicts a probability of 0.5%. 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 previous section establishes that accounting for social groups is important for reconciling the theory of voting with the empirical evidence. This raises a number of questions: how large is the population in organized groups relative to independents? what is the nature of these social groups?

With respect to the first question, recall that the model does not identify α_s and β_s independently, that is, the share of the population organized on side s and the probability that groups on that side vote for their preferred alternative in equilibrium. Instead, it identifies only $\theta_s = \alpha_s\beta_s$. The model therefore provides only a set estimate for α_s: since $\beta_s \in [0, 1]$, we can conclude that $\alpha_s \in [\theta_s, 1]$, so the estimated value $\widehat{\theta}_s$ is a lower bound on the share of the population organized on side s. Figure 3 plots the distribution of these side-specific lower bounds by 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 cities, for instance, for each side about half of the observations 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 cities. For each side, more than 80% 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 15%. 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

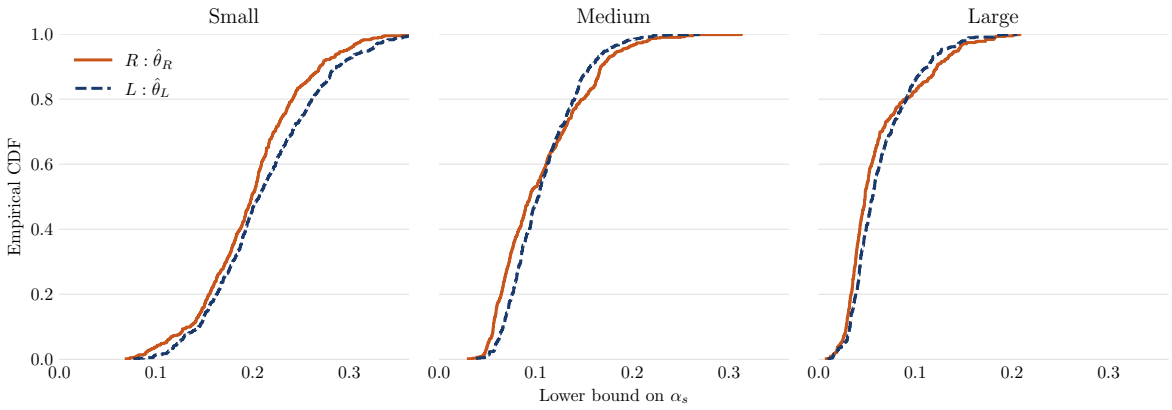


Figure 3: Distribution of the lower bound on the organized share of the electorate, $\alpha_s \geq \theta_s$, by electorate-size stratum (small, medium, and large jurisdictions). Each panel reports the empirical distribution of the election-level bound for R and L.

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.

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 59.1, with a median of 51. The estimated size, however, varies with the size of the electoral district. In small jurisdictions, the mean estimated group size is 19.5; it is 51.3 in medium jurisdictions; and it is on average 118.2 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 within each electorate-size stratum.[38] Groups are always small relative to the electorate: the estimated share exceeds 0.05 of eligible voters in only about a tenth of elections and never exceeds 0.20, and its full-sample median is 0.012. Moreover,

[38]It's worth noting that the CDFs for the three size strata only barely overlap – if the observed distribution of results could be rationalized by a smaller number of voters than the turnout, we would expect them to be effectively the same. Instead, the group sizes are meaningfully different even as proportions of n as n grows.

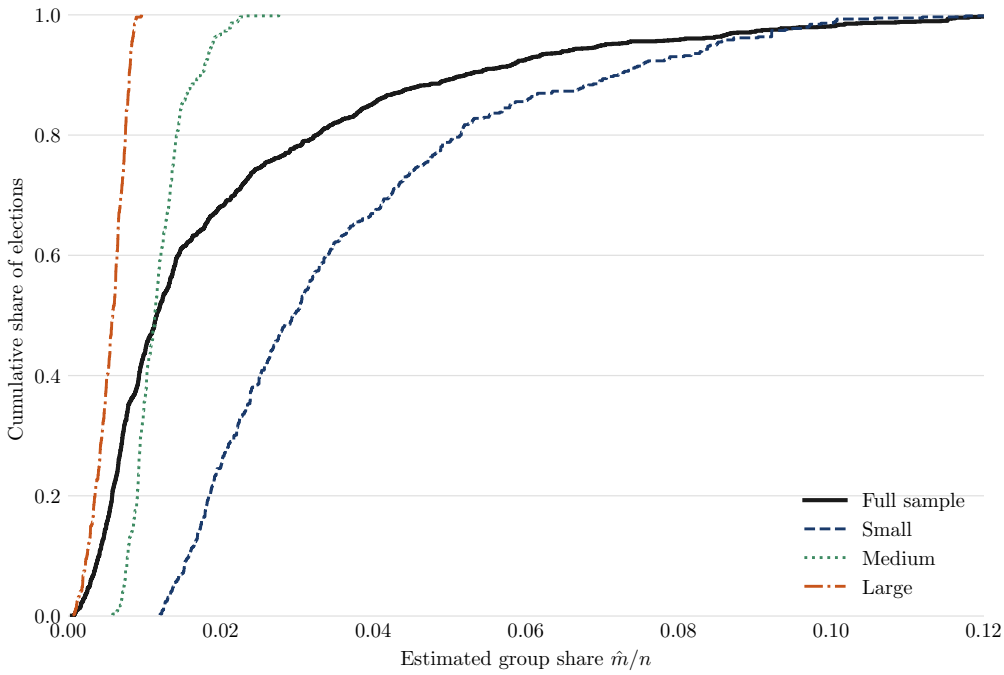


Figure 4: Empirical CDF of the estimated group share $\hat{m}/n$, for the full sample and within each electorate-size stratum.

although m rises with the size of the jurisdiction, the estimated *share* falls, averaging 0.036 in small jurisdictions, 0.012 in medium ones, and 0.006 in large ones. In absolute terms, m rarely becomes large: 93% of districts are associated with an m lower than 120. Even in the largest jurisdictions in the sample, the Dallas County justice precincts, with more than 770,000 eligible voters, the estimated group size is around 400—less than a tenth of a 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 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.

Despite being highly stylized, our model also provides some insights into the nature of the groups. As shown in Table 7, we find three factors that determine group size. First, group size is significantly correlated with both the Baptist population share and the average size of Southern Baptist congregations in the county. This is natural, since Baptist churches are organizations that are likely to be highly interested in liquor referenda. This is also confirmed by the fact that turnout is positively correlated with the presence of Baptist churches, and that Baptist presence is associated with higher turnout against liquor liberalization. The model also highlights that groups are larger

in city-wide elections. This seems consistent with the hypothesis that cities may be associated with more structured and better organized groups, such as unions or other nonprofit organizations.

5.2.3 Comparison to other voting models

As noted above, the restricted model with $m = 1$ and the more general model with groups can be easily ranked since they are nested: this led us to reject the hypothesis that $m = 1$ in our model. Here we compare our model with groups with three alternative models that are not nested, but have been or can be estimated on the same data we use. These models, therefore, can be compared to ours since they generate a likelihood function on the same event, that is the vote cast in the liquor referenda. The models are the following:[39]

- **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

[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 directly to single-district elections, and reports estimates of a referendum adaptation in which equilibrium effort is proxied by its determinants (Table 21).

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 is the model's own predictive distribution of outcomes. 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. The likelihood ranking is unambiguous. The grouped model attains a count log likelihood of −36,814, compared with −1,023,634 for the $m = 1$ benchmark, −1,301,477 for CCM, −1,260,610 for the group rule-utilitarian model, and −1,328,786 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 $\Pr(|M| > 0.10) = 0.010$, the $m = 1$ benchmark predicts 0.010, and the expressive model predicts 0.084, all against an empirical frequency of 0.209. 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 $\Pr(|M| > 0.05) = 0.378$, versus an empirical value of 0.402; the corresponding value in the grouped model is 0.371. It also predicts $\Pr(|M| > 0.10) = 0.160$, versus an empirical value of 0.209; the corresponding value in the grouped model is 0.168.

Still, the group rule-utilitarian model performs much worse than the grouped model in terms of the likelihood function: −1,260,610 versus −36,814. The reason is clear from the way the model

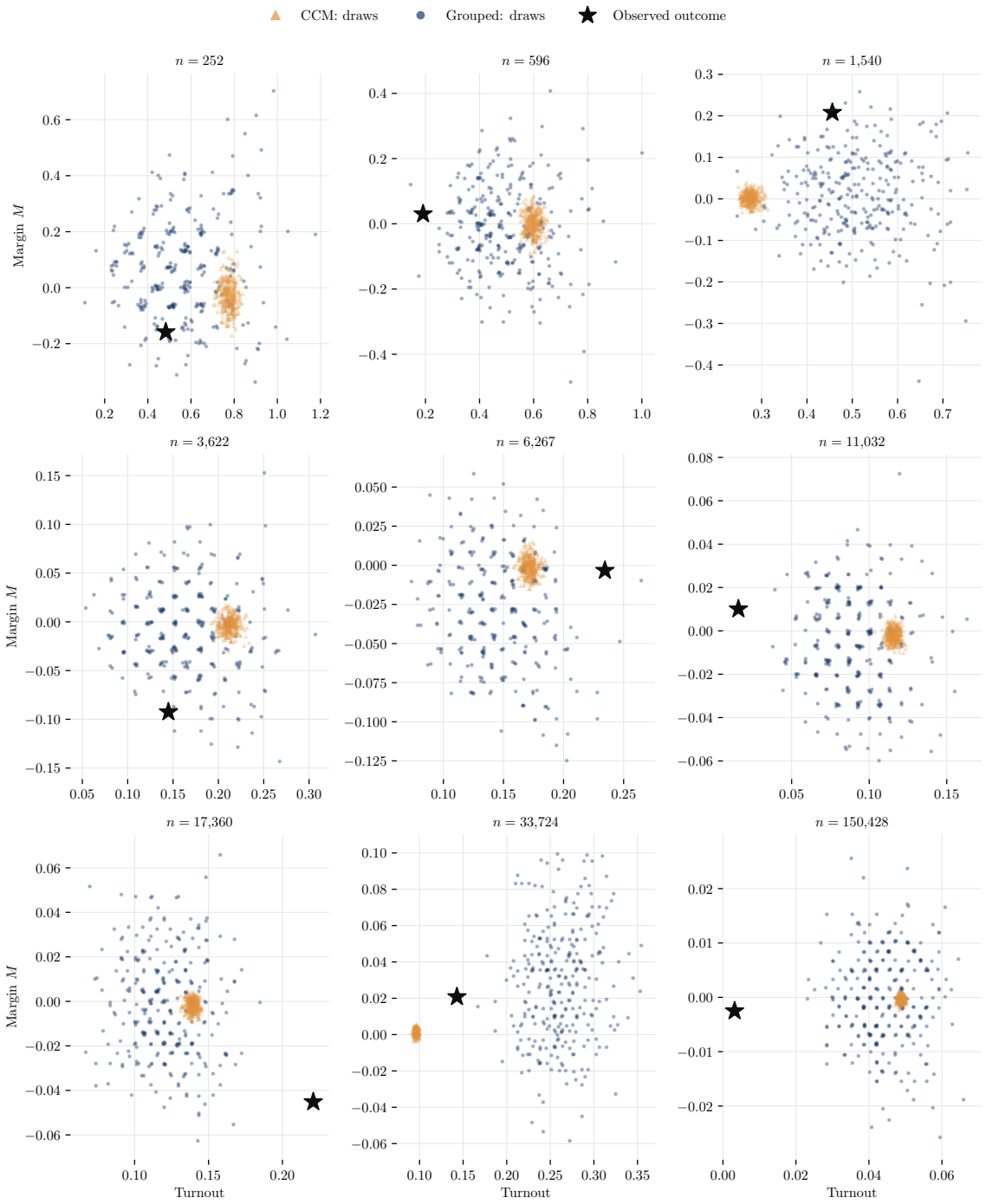


Figure 5: Predictive distributions election by election under the CCM pivotal-voter model, for nine elections at fixed quantiles of the electorate size. Red points: 300 simulated draws from the CCM model; blue points: 300 draws from the grouped model; star: the observed outcome.

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

$$\big(T_i(\mu_i), M_i(\mu_i)\big) = \big(\mu_i\sigma_{R,i} + (1-\mu_i)\sigma_{L,i},\ \mu_i\sigma_{R,i} - (1-\mu_i)\sigma_{L,i}\big). \tag{24}$$

The locus of possible outcomes generated by the model with a continuum of voters is the straight line segment with slope $-\frac{\sigma_{L,i}+\sigma_{R,i}}{\sigma_{L,i}-\sigma_{R,i}}$ in the (T, M) space, defined by equation (24) as μ_i varies (corresponding to different realizations). In the discrete-voter version we estimate, 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: Alternative Voting-Behavior Models

Statistic	Data	GRU	Expressive	CCM	$m = 1$	Grouped
Count log likelihood	–	-1,260,610	-1,328,786	-1,301,477	-1,023,634	-36,814
Turnout	0.247	0.191	0.177	0.297	0.278	0.268
SD turnout	0.190	0.118	0.092	0.200	0.171	0.184
SD margin	0.096	0.073	0.051	0.023	0.033	0.090
$\Pr(\lvert M\rvert > 0.05)$	0.402	0.378	0.251	0.049	0.113	0.371
$\Pr(\lvert M\rvert > 0.10)$	0.209	0.160	0.084	0.010	0.010	0.168
$\Pr(\lvert M\rvert > 0.15)$	0.114	0.056	0.015	0.002	0.001	0.087
$\Pr(\lvert M\rvert > 0.20)$	0.062	0.018	0.002	0.000	0.000	0.047
Vuong z vs. grouped	–	6.99 ($p < 0.001$)	6.74 ($p < 0.001$)	6.77 ($p < 0.001$)	7.44 ($p < 0.001$)	–
Clarke share vs. grouped	–	0.868 ($p < 0.001$)	0.943 ($p < 0.001$)	0.974 ($p < 0.001$)	0.945 ($p < 0.001$)	–

Notes: Full sample of 1,703 elections. The GRU column replicates the group rule-utilitarian model of Coate and Conlin (2004); the expressive and CCM columns replicate Coate et al. (2008). Model entries are simulation moments; the count log likelihood evaluates each model on the observed vote totals. The GRU column integrates over the fitted support-share distribution by quadrature. Vuong and Clarke statistics compare each column with the grouped model on the per-election log likelihoods (Vuong clustered by county); positive values and shares above one half favor the grouped model. $M = (V_R - V_L)/n$. Estimation and evaluation details are in Appendix II.

[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 $\mu\sigma_s - (1-\mu)\sigma_o$ with probability one.

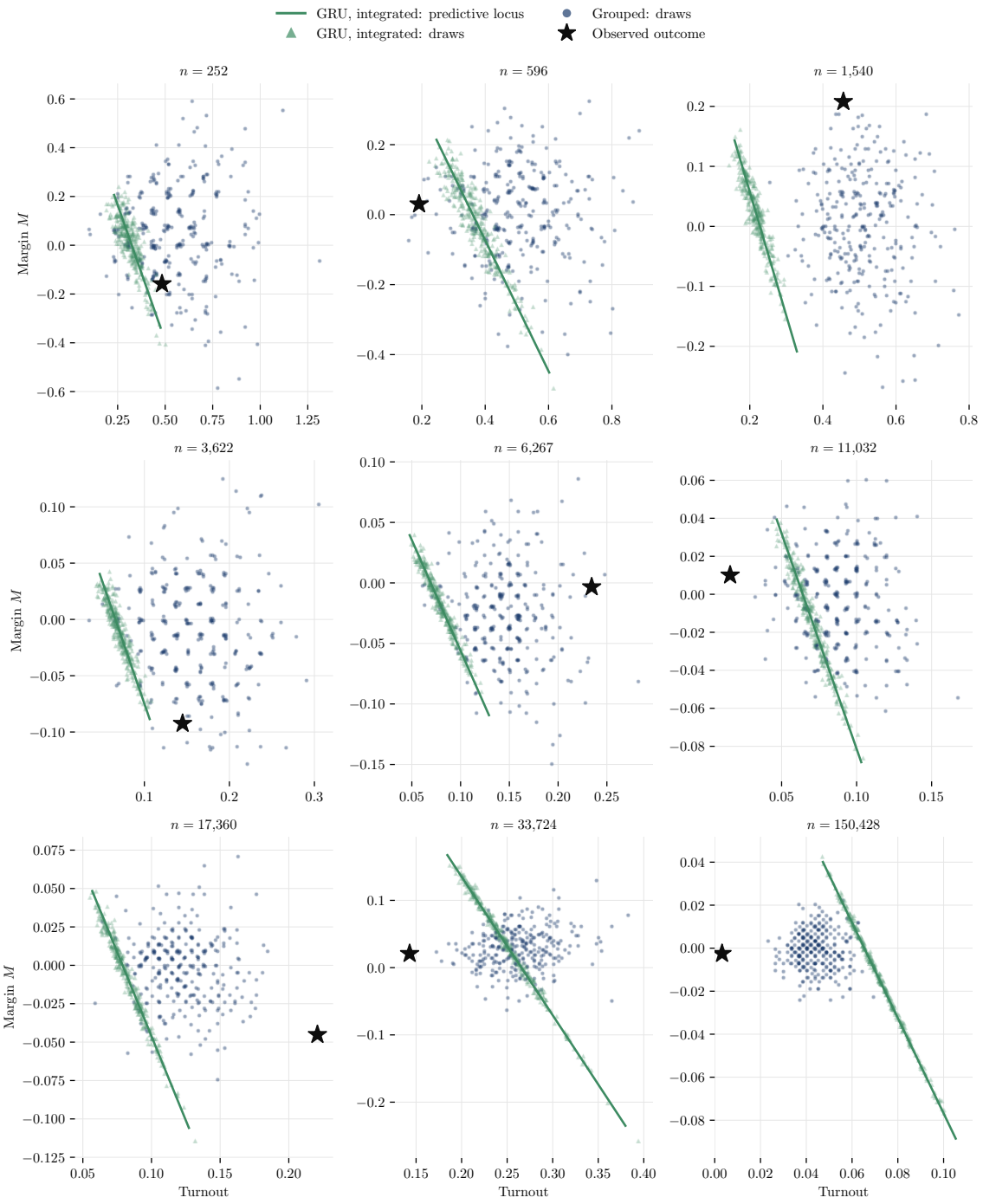


Figure 6: Predictive distributions election by election, for nine elections at fixed quantiles of the electorate size. Green line and points: the group rule-utilitarian model's predictive locus and simulated draws; blue points: the grouped model's simulated draws; star: the observed outcome. As the electorate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

Since all models considered here generate probability distributions over the same set of observable outcomes, namely realized vote totals in each district, their log likelihoods are directly comparable. We first use the Vuong test to rank the models. The bottom rows of Table 8 report the Vuong (1989) model-selection statistic z for each alternative M against the grouped model G.[41] Positive values of z support our model with groups: the statistic is asymptotically $N(0,1)$, so values above 2.56 support the grouped model at the 1% level. The clustered statistics range from 6.74 to 7.44 across the alternatives.

Table 8 adds a complementary comparison: the sign test of Clarke (2007), a distribution-free alternative to the Vuong statistic: it asks in what share of elections the grouped model assigns a higher likelihood to the observed outcome than the alternative, without weighting elections by the size of the log-likelihood difference. The grouped model fits better in 87 to 97 percent of elections, and the binomial null of one half is rejected at any conventional level against every alternative.

5.3 Robustness of the empirical results

The results of this section are robust along several dimensions; here we summarize the main exercises, which are reported in full in Appendix II. First, although the sample spans more than eight decades, the findings are not an artifact of pooling structurally different eras: re-estimating the full grouped model separately for elections held before 1970, in 1970–1989, and in 1990 or later, with the leakage parameter κ re-profiled within each period, yields the same qualitative conclusions (Appendix Table 11). Second, the results do not hinge on the auxiliary functional-form choices of Section 4.2: the grouped model's advantage over the $m = 1$ benchmark exceeds 970,000 log-likelihood points at every value of κ on the profiling grid (Appendix Figure 9), and it survives replacing the baseline electorate-size scaling of the residual individual-vote component with alternative specifications (Appendix Table 15). Third, the fitted valuation ratios $\hat{r}$ of equation (18) are tightly clustered around one—every election lies between 0.76 and 1.24—consistent with the assumptions of Propositions 2-4. The appendix also reports the grouped-model estimates within each electorate-size stratum, the coefficient estimates of the $m = 1$ benchmark, estimates with an extended covariate set and on the subsample of first referenda, and additional diagnostics for the alternative voting models of Section 5.2.3.

[41] Let $d_i = \ell_i^{(G)} - \ell_i^{(M)}$ be the per-election log-likelihood difference. Under the null hypothesis that the two models are equally close to the true data-generating process, $\sum_i d_i$, scaled by its standard deviation, is asymptotically standard normal. Because elections within a county share organizational infrastructure and repeat over time, we compute the variance clustered at the county level: $z = \left(\sum_i d_i\right) / \sqrt{\sum_c \left(\sum_{i \in c} (d_i - \bar{d})\right)^2}$, where c indexes the 217 counties and $\bar{d}$ is the sample mean of d_i.

5.4 Out of sample validation

The model's performance is validated out of sample. We randomly hold out 10 percent of the elections, re-estimate both the $m = 1$ benchmark and the grouped model on the remaining 90 percent, and score the fitted models on the held-out elections. The two models predict held-out turnout comparably well, but only the grouped model reproduces the dispersion of the margin of victory: margins of victory larger than 10 percent occur in 21.8% of the held-out elections and in 17.7% of the grouped model's simulations, against 1.2% under $m = 1$. The failure of the pivotal-voter model in the margin dimension is thus not an artifact of in-sample flexibility—it persists out of sample—and the ordering of the models is unchanged across ten independent 90/10 splits (Appendix Tables 16–19).

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 $\frac{1}{2}Piv_0^L(m_L, m_R)$, 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: $\sum_{l=1}^{m_L-1} Piv_{-l}^L(m_L, m_R) + \frac{1}{2}Piv_{-m_L}^L(m_L, m_R)$, 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:

$$c = v_L \left(\begin{array}{c} \frac{1}{2}Piv_0^L(m_L, m_R) \\ + \sum_{l=1}^{m_L-1} Piv_{-l}^L(m_L, m_R) + \frac{1}{2}Piv_{-m_L}^L(m_L, m_R) \end{array} \right) = v_L Piv_L. \tag{25}$$

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 (25) 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 $Piv_{-1}^L(m_L, m_R)$ 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 $Piv^I = \frac{1}{2}Piv_0^L(m_L, m_R) + \frac{1}{2}Piv_{-1}^L(m_L, m_R)$: 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 $Piv_{-1}^L(m_L, m_R) > 0$ when only groups vote, but, as mentioned above, $Piv_{-1}^L(m_L, m_R)$ may be strictly positive when $m_R > m_L$. Still, we have:

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

where we stress that the last inequality is strict. To see this it is sufficient to note that even if both $\frac{1}{2}Piv_{-1}^L(m_L, m_R)$ and $\sum_{l=2}^{m_L-1} Piv_{-l}^L(m_L, m_R)$ are zero, we always must have $Piv_{-m_L}^L(m_L, m_R) > 0$. 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, the key observation is 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. Using Myerson approximation, 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 prefer to be 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. Without i, the group becomes unable to create a tie in the event in which j is m votes below the other candidate. After the defection, the group is pivotal only if the other groups tie, i.e., in the same event as independents. In the equilibrium of Proposition 2, groups are indifferent between voting or abstaining as a block: it follows that the group finds it optimal to abstain after the defection. 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 both types of groups mix with interior probability, we have:

$$c = v_j \mathcal{P}^j(\beta) \simeq \frac{v_j e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2 \frac{n}{m}}}{4\left(\pi \frac{n}{m}\sqrt{\alpha_R\alpha_L\beta_R\beta_L}\right)^{\frac{1}{2}}}\left[1+\left(\frac{\alpha_{-j}\beta_{-j}}{\alpha_j\beta_j}\right)^{1/2}\right] \tag{26}$$

where the right-hand side is the group's expected instrumental benefit from voting when it behaves as a block. For simplicity, we approximate the right-hand side using Myerson's formula and, here and below, ignore the approximation error. 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 (26). Under this approximation, it follows that the net 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 (26) 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. This implies that i is willing to obey only if:

$$\frac{v_j e^{-\left(\sqrt{\alpha_R\beta_R}-\sqrt{\alpha_L\beta_L}\right)^2 \frac{n}{m}}}{4\left(\pi \frac{n}{m}\sqrt{\alpha_R\alpha_L\beta_R\beta_L}\right)^{\frac{1}{2}}}\left(\frac{\alpha_{-j}\beta_{-j}}{\alpha_j\beta_j}\right)^{1/2}+\pi^j \geq c \tag{27}$$

Using (5), we can approximately rewrite the first term on the left-hand side of (27) as $cv_{-j}/(v_R+v_L)$, and obtain the following lower bound for π^j:

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

Under the Myerson approximation, this 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, there exists an equilibrium, and in some cases exclusively equilibria, 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. 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 the empirical evidence. In particular, it can account for a discrepancy repeatedly emphasized in the literature: existing models have considerable difficulty explaining why elections produce large margins of victory with significant probability.

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. 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:

$$\frac{c}{v_L} > \frac{c}{v_I} \geq \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\beta_L\alpha_L n}(\beta_L\alpha_L n)^k}{k!}\right)\left(\frac{e^{-\beta_R\alpha_R n}(\beta_R\alpha_R n)^k}{k!}\right)\left(1+\frac{\beta_R\alpha_R n}{k+1}\right) \tag{28}$$

since $v_L < v_I$. It follows from (28) 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 (28). 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:

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

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: 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. 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. First, observe that in a large election we can assume that R-types vote with positive probability. Indeed, if this were not the case then only independents would vote and we would have $\frac{c_t}{v_I} = \frac{1}{2} e^{-\sigma_t \alpha_I n_t}$, therefore: $T_t^L = \log\left(\frac{v_I}{2c_t}\right) \to \infty$ and the result would be proven.

We next prove that $\sigma_t \alpha_I n_t > \alpha_R \beta_{R,t} n_t$. Since $\sigma_t = 0$ is impossible, we have $\sigma_t \in (0, 1)$. Moreover, the argument in Lemma 1 then implies $\beta_{L,t} = 0$. We must therefore have:

$$\frac{c_t}{v_I} = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\sigma_t\alpha_I n_t}(\sigma_t\alpha_I n_t)^k}{k!}\right)\left(\frac{e^{-\alpha_R\beta_{R,t}n_t}(\alpha_R\beta_{R,t}n_t)^k}{k!}\right)\left(1 + \frac{\alpha_R\beta_{R,t}n_t}{k+1}\right) \tag{29}$$

$$\frac{c_t}{v_R} \leq \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\sigma_t\alpha_I n_t}(\sigma_t\alpha_I n_t)^k}{k!}\right)\left(\frac{e^{-\alpha_R\beta_{R,t}n_t}(\alpha_R\beta_{R,t}n_t)^k}{k!}\right)\left(1 + \frac{\sigma_t\alpha_I n_t}{k+1}\right) \tag{30}$$

where the right-hand side of equation (29) 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 (30) 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 we suppose to the contrary that $\sigma_t\alpha_I n_t \leq \alpha_R\beta_{R,t}n_t$, then (29) and (30) imply:

$$\begin{aligned}\frac{c_t}{v_I} &= \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\sigma_t\alpha_I n_t}(\sigma_t\alpha_I n_t)^k}{k!}\right)\left(\frac{e^{-\alpha_R\beta_{R,t}n_t}(\alpha_R\beta_{R,t}n_t)^k}{k!}\right)\left(1 + \frac{\alpha_R\beta_{R,t}n_t}{k+1}\right)\\ &\geq \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\sigma_t\alpha_I n_t}(\sigma_t\alpha_I n_t)^k}{k!}\right)\left(\frac{e^{-\alpha_R\beta_{R,t}n_t}(\alpha_R\beta_{R,t}n_t)^k}{k!}\right)\left(1 + \frac{\sigma_t\alpha_I n_t}{k+1}\right) \geq \frac{c_t}{v_R},\end{aligned}$$

a contradiction, since $v_I > v_R$. Hence $\sigma_t\alpha_I n_t > \alpha_R\beta_{R,t}n_t$.[42]

We can now prove the statement of Lemma 1.1. Suppose to the contrary that there is a sequence $(c_t, n_t, \sigma_t, \beta_{R,t})_{t=1}^{\infty}$ such that $c_t \to 0$, $n_t \to \infty$ and $\sigma_t\alpha_I n_t = T_t^L \nrightarrow \infty$. Then there is a subsequence

[42] This implies that the probability L wins is greater than $1/2$.

along which T_t^L is bounded. Passing to a further subsequence, assume $T_t^L \to T_\infty^L < \infty$. 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 (29), we have:

$$\begin{aligned} 0 &= \lim_{t\to\infty} \frac{c_t}{v_I} = \frac{1}{2} \lim_{t\to\infty} \sum_{k=0}^{\infty} \left(\frac{e^{-\sigma_t \alpha_I n_t} (\sigma_t \alpha_I n_t)^k}{k!} \right) \left(\frac{e^{-\alpha_R \beta_{R,t} n_t} (\alpha_R \beta_{R,t} n_t)^k}{k!} \right) \left(1 + \frac{\alpha_R \beta_{R,t} n_t}{k+1} \right) \\ &\geq \frac{1}{2} \lim_{t\to\infty} \left[(1 + \alpha_R \beta_{R,t} n_t)\, e^{-\left(\alpha_I \sigma_t + \alpha_R \beta_{R,t}\right) n_t} \right] \\ &= \frac{1}{2} \left(1 + T_\infty^R\right) e^{-\left(T_\infty^L + T_\infty^R\right)} > 0, \end{aligned}$$

which is a contradiction. We conclude that $\sigma_t n_t \to \infty$ as $t \to \infty$. ■

Using Lemma 1.1, we can now prove that for any sequence c_t, n_t and associated strategy $\sigma_t, \beta_{R,t}$ such that $c_t \to 0$, $n_t \to \infty$, the probability that L wins converges to 1, that is, $P(\sigma_t, \beta_{R,t}) \to 1$. Note that expected turnout for L and R are, respectively, $\sigma_t \alpha_I n_t$ and $\alpha_R \beta_{R,t} n_t$. If $\alpha_R \beta_{R,t} n_t \to T_\infty^R < \infty$, then L's margin of victory diverges at infinity, so L wins with arbitrarily high probability as $t \to \infty$, which proves the result. Assume instead that $\alpha_R \beta_{R,t} n_t \to \infty$ as well. There are two possibilities: (1) there is a subsequence $t \to \infty$ such that $\beta_{R,t} = 1$; or (2) there for all sequences $t \to \infty$, $\beta_{R,t} < 1$. In either case, using the limit formulas in Myerson (2000) we must have that as $t \to \infty$:

$$c \leq v_R \frac{e^{-\left(\sqrt{\beta_{R,t}\alpha_R n_t} - \sqrt{\sigma_t \alpha_I n_t}\right)^2}}{4\left(\pi \sqrt{\beta_{R,t}\alpha_R n_t \cdot \sigma_t \alpha_I n_t}\right)^{\frac{1}{2}}} \left(\frac{\left(\beta_{R,t}\alpha_R n_t\right)^{\frac{1}{2}} + \left(\sigma_t \alpha_I n_t\right)^{\frac{1}{2}}}{\left(\beta_{R,t}\alpha_R n_t\right)^{\frac{1}{2}}} \right) (1 + o(1)) \tag{31}$$

$$c = v_I \frac{e^{-\left(\sqrt{\beta_{R,t}\alpha_R n_t} - \sqrt{\sigma_t \alpha_I n_t}\right)^2}}{4\left(\pi \sqrt{\beta_{R,t}\alpha_R n_t \cdot \sigma_t \alpha_I n_t}\right)^{\frac{1}{2}}} \left(\frac{\left(\beta_{R,t}\alpha_R n_t\right)^{\frac{1}{2}} + \left(\sigma_t \alpha_I n_t\right)^{\frac{1}{2}}}{\left(\sigma_t \alpha_I n_t\right)^{\frac{1}{2}}} \right) (1 + o(1)) \tag{32}$$

Note that that the limit formulas are exact as $t \to \infty$ since turnout for both L and R diverges at infinity. For this equilibrium to exist, we need to have: $\frac{v_I}{v_R} \leq \frac{\left(\sigma_t \alpha_I n_t\right)^{\frac{1}{2}}}{\left(\beta_{R,t}\alpha_R n_t\right)^{\frac{1}{2}}} (1 + o(1))$. This implies that expected turnout for L is $T_t^L = \sigma_t \alpha_I n_t \geq \left(\frac{v_I}{v_R}\right)^2 \beta_{R,t}\alpha_R n_t$ (omitting for simplicity the vanishing term), and expected turnout for R is ${T_t}^R = \beta_{R,t}\alpha_R n_t$. In particular, the margin of victory for L is at least $\left[\left(\frac{v_I}{v_R}\right)^2 - 1 \right] \beta_{R,t}\alpha_R n_t$, so it diverges at infinity. As $t \to \infty$, the law of large numbers implies that L wins with arbitrarily high probability.

Since for any sequence c_t, n_t and associated strategy $\sigma_t, \beta_{R,t}$ such that $c_t \to 0$, $n_t \to \infty$, we have $\lim_{t\to\infty} P^L(\sigma_t, \beta_{R,t}) \to 1$, we must have 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 note, suppose the proposition were false, so $\lim_{c\to 0} \liminf_{n\to\infty} P_n^L(c) \nrightarrow 1$. 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 $\liminf$, for each t we could then 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 $1 > \beta_R > \frac{\alpha_L}{\alpha_R}\beta_L > 0 = \sigma$ and the probability that L wins is strictly less than $1/2$.

Proof: We proceed in four steps. In Step 1, assuming that independents abstain, we establish preliminary properties of equilibria of the modified game played by the R- and L-groups. In Step 2, we show that equilibrium participation satisfies $\beta_R^n > \left(\frac{\alpha_L}{\alpha_R}\right)\beta_L^n$ and that $\beta_R^n \in (0,1)$. In Step 3, we show that, for sufficiently large n, $\beta_L^n \in (0,1)$ as well. Finally, in Step 4, 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. Assume the voting probability for independents is zero, $\sigma_n = 0$, and consider the modified game played by groups R and groups L in which groups can either abstain or vote as a group with probabilities β_j^n for $j = L, R$. By a standard argument, this game must have a mixed equilibrium $\beta_R^n \in [0,1]$, $\beta_L^n \in [0,1]$. We now make three observations. First, we cannot have an equilibrium with $\beta_R^n = \beta_L^n = 0$, since $c < (1/2)v_L$. Second, there must be a $\overline{n}$ such that for $n > \overline{n}$, $\beta_R^n = \beta_L^n = 1$ cannot be an equilibrium, since the pivot probability would otherwise converge to zero for both groups, making it suboptimal to vote. Third, we prove that we cannot have $\beta_R^n = 1$ and $\beta_L^n > 0$. There are two cases: (a) $\beta_R^n = 1$ and $\beta_L^n \to \beta_L \in (0,1]$, so $\beta_L^n n \to \infty$; and (b) $\beta_R^n = 1$ and $\lim_{n\to\infty} \beta_L^n = 0$. Suppose first that $\beta_R^n = 1$ and $\beta_L^n \to \beta_L \in (0,1]$. In this case for n large, turnout for both L and R diverges at infinity, so we can apply the Myerson (2000) approximation formula and write:

$$\frac{c}{v_L} > \lim_{n\to\infty} \frac{e^{-\left(\sqrt{\alpha_R}-\sqrt{\alpha_L\beta_L^n}\right)^2 \frac{n}{m}}}{4\left(\frac{\pi}{m}\sqrt{\alpha_R\alpha_L\beta_L^n}\cdot n\right)^{\frac{1}{2}}}\left(1+\frac{1}{\sqrt{\alpha_L\beta_L^n}}\right) \to \frac{e^{-\left(\sqrt{\alpha_R}-\sqrt{\alpha_L\beta_L}\right)^2 \frac{n}{m}}}{4\left(\frac{\pi}{m}\sqrt{\alpha_R\alpha_L\beta_L}\cdot n\right)^{\frac{1}{2}}}\left(1+\frac{1}{\sqrt{\alpha_L\beta_L}}\right) \to 0,$$

a contradiction.

Suppose instead that $\beta_R^n = 1$ and $\beta_L^n > 0$ but $\lim_{n\to\infty} \beta_L^n = 0$. Then for n large, the expected margin of victory for R converges to $MV_n^R = (\alpha_R - \alpha_L\beta_L^n)\, n$, which diverges at infinity. This implies that the pivot probability for L converges to zero. As $n \to \infty$, L's pivot probability must become strictly lower than $\frac{c}{v_L}$, thus implying $\beta_L^n = 0$ for sufficiently large values of n: a contradiction.

Step 2. We now prove that we cannot have $\beta_L^n \geq \frac{\alpha_R}{\alpha_L}\beta_R^n$. If this were the case, by the first two observations, we would have either $\beta_L^n \geq \frac{\alpha_R}{\alpha_L}\beta_R^n > 0$ or $\beta_L^n > \frac{\alpha_R}{\alpha_L}\beta_R^n = 0$. From Step 1 we have

$\beta_R^n \in [0,1)$. It follows that, in both cases, we must have:

$$\begin{aligned} \frac{c}{v_R} &\geq \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^n n}(\frac{\alpha_R}{m}\beta_R^n n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^n n}(\frac{\alpha_L}{m}\beta_L^n n)^k}{k!}\right)\left(1+\frac{\frac{\alpha_L}{m}\beta_L^n n}{k+1}\right) \qquad (33)\\ &\geq \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^n n}(\frac{\alpha_R}{m}\beta_R^n n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^n n}(\frac{\alpha_L}{m}\beta_L^n n)^k}{k!}\right)\left(1+\frac{\frac{\alpha_R}{m}\beta_R^n n}{k+1}\right) \geq \frac{c}{v_L} \end{aligned}$$

a contradiction since $v_R > v_L$. In (33), the right-hand side of the first inequality is the pivot probability for a group of type R: the probability that groups are tied, or the groups for R are 1 group short. It follows that we must have an equilibrium in which $\beta_R^n > \frac{\alpha_L}{\alpha_R}\beta_L^n \geq 0$, so $\beta_R^n \in (0,1)$.

Step 3. We now prove that there is a $\overline{n}$ such that for $n > \overline{n}$ we have $\beta_R^n \in (0,1)$, $\beta_L^n \in (0,1)$. Assume by contradiction that this is not true. A possibility is that there is a sequence of equilibria in which $\beta_L^n = 0$, $\beta_R^n \in (0,1]$ as $n \to \infty$. Since $\beta_L^n = 0$, we can trivially rule out $\beta_R^n = 1$ for large n, so $\beta_R^n \in (0,1)$, implying the equilibrium condition for R groups: $\frac{c}{v_R} = \frac{1}{2}e^{-\frac{\alpha_R}{m}\beta_R^n n} \Leftrightarrow \frac{\alpha_R}{m}\beta_R^n n = \log\left(\frac{v_R}{2c}\right)$, and $\beta_L^n = 0$ implies the equilibrium inequality condition for L groups:

$$\begin{aligned} \frac{c}{v_L} &\geq \frac{1}{2}e^{-\frac{\alpha_R}{m}\beta_R^n n}\left(1+\frac{\alpha_R}{m}\beta_R^n n\right) \Leftrightarrow \frac{v_R}{v_L} \geq 1+\log\left(\frac{v_R}{2c}\right)\\ &\Leftrightarrow e^{\frac{v_R}{v_L}-1} \geq \frac{v_R}{2c} \Leftrightarrow c \geq \bar{c}, \end{aligned}$$

a contradiction.

The last case to consider is when $\beta_L^n = 1$. In this case $\beta_R^n > \frac{\alpha_L}{\alpha_R} > 0$, so turnout for both groups diverges at infinity for both groups. In this case, since $\alpha_R\beta_R^n > \alpha_L\beta_L^n$, the pivot probability for both groups converges to zero, implying a contradiction with the assumption that $\beta_L^n > 0$ and $\beta_R^n > 0$.

Step 4. Consider a sequence of modified voting games in which $\sigma_n = 0$, and an associated sequence of equilibria for the groups $\beta_L^n \in (0,1)$, $\beta_R^n \in (0,1)$. We now prove that there is a $\Delta^* > 0$ such that if $v_I \leq v_R(1+\Delta^*)$, then for a sufficiently large n it is indeed optimal for the independents to abstain if they expect the groups' strategies to be β_L^n, β_R^n.

The equilibrium conditions of the modified game for a mixed equilibrium β_L^n, β_R^n are:

$$\frac{c}{v_R} = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^n n}(\frac{\alpha_R}{m}\beta_R^n n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^n n}(\frac{\alpha_L}{m}\beta_L^n n)^k}{k!}\right)\left(1+\frac{\frac{\alpha_L}{m}\beta_L^n n}{k+1}\right) = \mathcal{P}_n^R(\beta) \qquad (34)$$

$$\frac{c}{v_L} = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^n n}(\frac{\alpha_L}{m}\beta_L^n n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^n n}(\frac{\alpha_R}{m}\beta_R^n n)^k}{k!}\right)\left(1+\frac{\frac{\alpha_R}{m}\beta_R^n n}{k+1}\right) = \mathcal{P}_n^L(\beta) \qquad (35)$$

where $\mathcal{P}_n^j(\beta)$ for $j = L, R$ is the pivotal probability for groups of type j. Subtracting (34) from (35), we obtain:

$$c\left[\frac{1}{v_L}-\frac{1}{v_R}\right] = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^n n}(\frac{\alpha_L}{m}\beta_L^n n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^n n}(\frac{\alpha_R}{m}\beta_R^n n)^k}{k!}\right)\frac{\frac{\alpha_R}{m}\beta_R^n n}{k+1}\left(1-\frac{\alpha_L\beta_L^n}{\alpha_R\beta_R^n}\right) \qquad (36)$$

Let $(\alpha_L\beta_L^n)/(\alpha_R\beta_R^n) = \xi^n$ and note that $\xi^n < \overline{\xi}$ for some $\overline{\xi} < 1$. If not, then there is a subsequence $\xi^{n_j} \to 1$ as $n_j \to \infty$. We would have:

$$\frac{c}{v_R} = \lim_{n_j\to\infty} \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^{n_j}n_j}(\frac{\alpha_R}{m}\beta_R^{n_j}n_j)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^{n_j}n_j}(\frac{\alpha_L}{m}\beta_L^{n_j}n_j)^k}{k!}\right)\left(1+\frac{\xi^{n_j}\frac{\alpha_R}{m}\beta_R^{n_j}n_j}{k+1}\right) = \frac{c}{v_L}, \tag{37}$$

a contradiction since $v_R > v_L$. Moreover, the next Lemma proves that $\xi^n > \underline{\xi}$ for some $\underline{\xi} > 0$.

Lemma 2.1. Suppose, by contradiction, that there is a subsequence along which $\xi_n \to 0$. From (37):

$$\frac{c}{v_R} = \lim_{n\to\infty} \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^{n}n}(\frac{\alpha_R}{m}\beta_R^{n}n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^{n}n}(\frac{\alpha_L}{m}\beta_L^{n}n)^k}{k!}\right) \tag{38}$$

Define $\lambda_j^n = \frac{\alpha_j\beta_j^n}{m}n$ for $j = L, R$. Equation (38) implies that λ_R^n cannot diverge. Indeed, if $\lambda_R^n \to \infty$ and $\frac{\lambda_L^n}{\lambda_R^n} \to 0$, then the probability on the right-hand side of (38) converges to zero, thus contradicting (38). Nor can $\lambda_R^n \to 0$. In that case, $\lambda_L^n \to 0$, so

$$\lim_{n\to\infty} \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^{n}n}(\frac{\alpha_R}{m}\beta_R^{n}n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^{n}n}(\frac{\alpha_L}{m}\beta_L^{n}n)^k}{k!}\right) \to \frac{1}{2},$$

contradicting $\frac{c}{v_R} < \frac{1}{2}$. Hence, after passing to a further subsequence, $\lambda_R^n \to \lambda_R \in (0,\infty)$, and $\lambda_L^n \to 0$.

The two limiting indifference conditions are then $\frac{c}{v_R} = \frac{1}{2}e^{-\lambda_R}$ and $\frac{c}{v_L} = \frac{1}{2}e^{-\lambda_R}(1+\lambda_R)$. Dividing the first by the second, gives $\lambda_R = \frac{v_R}{v_L} - 1$. Substituting this into the first gives $c = \frac{v_R}{2}e^{1-v_R/v_L} = \bar{c}$, contradicting the maintained assumption $c < \bar{c}$. It follows that $\liminf_{n\to\infty}\xi_n > 0$. This proves the required uniform lower bound. ■

We can write:

$$\frac{c}{1-\xi_n}\left[\frac{1}{v_L}-\frac{1}{v_R}\right] = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^{n}n}(\frac{\alpha_L}{m}\beta_L^{n}n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^{n}n}(\frac{\alpha_R}{m}\beta_R^{n}n)^k}{k!}\right)\frac{\frac{\alpha_R}{m}\beta_R^n n}{k+1} \tag{39}$$

Independents find it optimal to abstain if:

$$\frac{c}{v_I} \geq \mathcal{P}_n^I(\beta) = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^{n}n}(\frac{\alpha_L}{m}\beta_L^{n}n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^{n}n}(\frac{\alpha_R}{m}\beta_R^{n}n)^k}{k!}\right) \tag{40}$$

where $\mathcal{P}^I(\beta)$ is the pivotal probability for independents. We have:

$$\begin{aligned}\mathcal{P}_n^I(\beta) &= \mathcal{P}_n^L(\beta) - \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^{-\frac{\alpha_L}{m}\beta_L^{n}n}(\frac{\alpha_L}{m}\beta_L^{n}n)^k}{k!}\right)\left(\frac{e^{-\frac{\alpha_R}{m}\beta_R^{n}n}(\frac{\alpha_R}{m}\beta_R^{n}n)^k}{k!}\right)\frac{\frac{\alpha_R}{m}\beta_R^n n}{k+1} \\ &\leq c\left[\frac{1}{v_L} - \frac{1}{1-\xi_n}\left[\frac{1}{v_L}-\frac{1}{v_R}\right]\right] = c\frac{v_L - \xi_n v_R}{(1-\xi_n)v_L v_R}\end{aligned} \tag{41}$$

where we used (35) and (39). From (40) and (41), we conclude that if:

$$\frac{c}{v_I} \geq c\frac{v_L - \xi_n v_R}{(1-\xi_n)\, v_L v_R} \Longleftrightarrow \frac{1}{v_I} \geq \frac{v_L - \xi_n v_R}{(1-\xi_n)\, v_L v_R} \tag{42}$$

then $\frac{c}{v_I} \geq c\frac{v_L - \xi_n v_R}{(1-\xi_n) v_L v_R} \geq \mathcal{P}_n^I(\beta)$. Since $\frac{1}{v_R} = \frac{v_L - \xi_n v_L}{(1-\xi_n) v_L v_R} > \frac{v_L - \xi_n v_R}{(1-\xi_n) v_L v_R}$, there is a $\Delta_n^* > 0$ such that:

$$\frac{1}{(1+\Delta_n^*)\, v_R} = \frac{(v_L - \xi_n v_R)}{(1-\xi_n)\, v_L v_R}$$

Moreover there must be a $\Delta^* > 0$ such that $\Delta_n^* \geq \Delta^*$. Assume not, then there is a subsequence $n_j \to \infty$ such that $\Delta_{n_j}^* \to 0$, but then:

$$\lim_{n_j \to \infty} \frac{1}{\left(1+\Delta_{n_j}^*\right) v_R} = \frac{1}{v_R} = \lim_{n_j \to \infty} \frac{\left(v_L - \xi_{n_j} v_R\right)}{\left(1-\xi_{n_j}\right) v_L v_R} < \lim_{n_j \to \infty} \frac{\left(v_L - \xi_{n_j} v_L\right)}{\left(1-\xi_{n_j}\right) v_L v_R} = \frac{1}{v_R},$$

where we note that $\frac{\left(v_L - \xi_{n_j} v_L\right)}{\left(1-\xi_{n_j}\right) v_L v_R} = \frac{1}{v_R}$ for any n: so we have a contradiction. It follows that $v_I \leq (1+\Delta^*)\, v_R$ implies:

$$\frac{1}{v_I} \geq \frac{1}{(1+\Delta^*)\, v_R} > \frac{1}{(1+\Delta_n^*)\, v_R} = \frac{(v_L - \xi_n v_R)}{(1-\xi_n)\, v_L v_R}$$

which in turn implies $\frac{c}{v_I} \geq \mathcal{P}_n^I(\beta)$. We conclude that for n sufficiently large, the independents find it optimal to abstain in the sequence of equilibria. ■

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: Consider equations (1)-(2) or equivalently (7)-(8). For c sufficiently small, there is a $n(c)$ such that for $n > n(c)$, they admit a unique solution $\psi_i^n(c) \in (0,1)$ for $i \in L, R$ such that: $\widehat{T}_i(c) = \frac{\alpha_i}{m} n \psi_i^n(c)$ is independent of n. These solutions are decreasing in c, and such that $\frac{\alpha_R \psi_R^n(c)}{\alpha_L \psi_L^n(c)} = (\frac{v_R}{v_L})^2 > 1$. Consider any sequence of these solutions corresponding to $c_l \to 0$ as $n_l \to \infty$ with $n_l \geq n(c_l)$. Along this sequence turnout diverges. From Myerson (2000), it follows that the approximate pivot probabilities appearing on the right-hand sides of (1)-(2) converge to the corresponding exact pivot probabilities of the modified voting game in which independents abstain, implying the existence of an exact equilibrium whose expected turnout is asymptotically close to the limiting solution. Convergence is guaranteed by the fact that the left-hand side of (7) is strictly decreasing in $\widehat{T}_i(c)$, so it has a unique intersection with the right-hand side. Details are presented in the online appendix. Specifically, the constructed equilibrium has the property that $\frac{T_R(c_l)}{T_L(c_l)} = \frac{\alpha_R \beta_R^n(c_l)}{\alpha_L \beta_L^n(c_l)} \to (\frac{v_R}{v_L})^2 > 1$ and

$T_R(c_l) \to \infty$. The argument in Section 3.1 shows that along the sequence of equilibria $\beta_R^n(c_l)$, $\beta_L^n(c_l)$ independents do not want to vote in a large economy if $v_I < v_L + v_R$.

Consider any sequence c_t, n_t such that $c_t \to 0$, $n_t \to \infty$ with $n_l \geq n(c_l)$, and an associated equilibrium sequence $\beta_{L,t}$, $\beta_{R,t}$ constructed as above. In correspondence to this sequence, we must have:

$$\frac{\mu_{MV}}{\sigma_{MV}} = \frac{T_R^n - T_L^n}{\sqrt{T_R^n + T_L^n}} = \frac{\sqrt{T_R^n}\left(1 - T_L^n/T_R^n\right)}{\sqrt{1 + T_L^n/T_R^n}} \to \sqrt{T_R^n}\frac{1 - \left(v_L/v_R\right)^2}{\sqrt{1 + \left(v_L/v_R\right)^2}} \to \infty$$

Chebyshev inequality implies that the probability R wins approaches 1 as $n \to \infty$. It follows that for c sufficiently small, there is a $n(c)$ and an equilibrium of the economies c, n with $n > n(c)$ in which independents abstain and L wins with probability lower than δ. ■

I.5 Proof of Proposition 4

Proposition 4: In a large election with organized groups, there is a ϑ^*, such that for $\vartheta < \vartheta^*$, $\sigma = 0$ in all equilibria.

Proof: Note that $\sigma_n = 1$ cannot be an equilibrium, otherwise L would win with arbitrarily high probability and independents would not find it optimal to vote. If there is an equilibrium in which the independents vote with positive probability, then it must be such that $\sigma_n \in (0,1)$. We proceed in two steps.

Step 1. First, we show that we must have either have $\beta_R^n \geq \frac{\alpha_L}{\alpha_R}\beta_L^n$, or $\beta_R^n = 1$ and thus $\beta_L^n \geq \alpha_R/\alpha_L$. If this condition is not satisfied, then $\frac{\alpha_L}{\alpha_R}\beta_L^n > \beta_R^n$ and $\beta_R^n < 1$. This implies $\beta_L^n \in (0,1]$. In this case, we would have:

$$\begin{aligned}
\frac{c}{v_R} &< \frac{c}{v_L} \leq \mathcal{P}_L\left(\sigma_n, \beta\right) \\
&< \sum\nolimits_{i=0}^{\infty} S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\left(\begin{array}{c} \frac{1}{2}\frac{(\sigma_n\alpha_I n)^{mi}e^{-\sigma_n\alpha_I n}}{(mi)!} + \sum_{j=mi+1}^{m(i+1)-1}\frac{(\sigma_n\alpha_I n)^{j}e^{-\sigma_n\alpha_I n}}{j!} \\ +\frac{1}{2}\frac{(\sigma_n\alpha_I n)^{m(i+1)}e^{-\sigma_n\alpha_I n}}{[m(i+1)]!}\end{array}\right) \\
&< \sum\nolimits_{i=0}^{\infty} S_{-i}\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\left(\begin{array}{c} \frac{1}{2}\frac{(\sigma_n\alpha_I n)^{mi}e^{-\sigma_n\alpha_I n}}{(mi)!} + \sum_{j=mi+1}^{m(i+1)-1}\frac{(\sigma_n\alpha_I n)^{j}e^{-\sigma_n\alpha_I n}}{j!} \\ +\frac{1}{2}\frac{(\sigma_n\alpha_I n)^{m(i+1)}e^{-\sigma_n\alpha_I n}}{[m(i+1)]!}\end{array}\right) = \mathcal{P}_R\left(\sigma_n, \beta\right)
\end{aligned}$$

since $\frac{\alpha_L\beta_L^n n}{m} > \frac{\alpha_R\beta_R^n n}{m}$ implies that $S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right) < S_{-i}\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)$. It follows that $\beta_R^n = 1$, a contradiction.

Step 2. Following Step 1, we have 2 cases: 1. $\sigma^n \in (0,1)$, $\beta_L^n \in [0,1)$, and $\beta_R^n \in [\frac{\alpha_L}{\alpha_R}\beta_L^n, 1]$; and 2. $\sigma_n \in (0,1)$, $\beta_R^n = 1$ and $\beta_L^n > \frac{\alpha_R}{\alpha_L}$ (since this case is possible only if $\beta_R^n < \frac{\alpha_L}{\alpha_R}\beta_L^n$, so $\frac{\alpha_R}{\alpha_L} < 1$).

Case 1: $\sigma_n \in (0,1)$, $\beta_L^n \in [0,1)$, and $\beta_R^n \in [\frac{\alpha_L}{\alpha_R}\beta_L^n, 1]$. The difference of two Poisson random variables x_1 and x_2 with means m_1 and m_2 is distributed according to the Skellam distribution

$S(m_1, m_2)$ with mean $m_1 - m_2$, and variance $m_1 + m_2$. We have:

$$\frac{c}{v_I} = \mathcal{P}_I(\sigma_n, \beta) = \frac{1}{2}\left[\begin{array}{c} S_0\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)e^{-\sigma_n\alpha_I n} \\ +\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\cdot\left(\frac{(\sigma_n\alpha_I n)^{mi-1}e^{-\sigma_n\alpha_I n}}{(mi-1)!} + \frac{(\sigma_n\alpha_I n)^{mi}e^{-\sigma_n\alpha_I n}}{(mi)!}\right)\end{array}\right] \tag{43}$$

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:

$$\begin{aligned}
&\frac{c}{v_L} \geq \mathcal{P}_L(\sigma_n, \beta) \\
&\geq \frac{1}{2}\left[\begin{array}{c} S_0\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)e^{-\sigma_n\alpha_I n} \\ +\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\left(\begin{array}{c}\frac{(\sigma_n\alpha_I n)^{mi-1}e^{-\sigma_n\alpha_I n}}{(mi-1)!} + \frac{(\sigma_n\alpha_I n)^{mi}e^{-\sigma_n\alpha_I n}}{(mi)!} \\ +\sum_{j=2}^{m}\frac{(\sigma_n\alpha_I n)^{(mi-j)}e^{-\sigma_n\alpha_I n}}{(mi-j)!}\end{array}\right)\end{array}\right] \\
&= \mathcal{P}_I(\sigma_n, \beta) + \frac{1}{2}\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\cdot\sum_{j=2}^{m}\frac{(\sigma_n\alpha_I n)^{(mi-j)}e^{-\sigma_n\alpha_I n}}{(mi-j)!} \\
&\Leftrightarrow \frac{c}{v_L} - \frac{c}{v_I} \geq \frac{1}{2}\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\cdot\sum_{j=2}^{m}\frac{(\sigma_n\alpha_I n)^{(mi-j)}e^{-\sigma_n\alpha_I n}}{(mi-j)!} = \chi^n
\end{aligned} \tag{44}$$

where in the top three lines we use the fact that $\mathcal{P}_L(\sigma_n, \beta)$ at least $\mathcal{P}_I(\sigma_n, \beta)$ plus the $1/2$ times the probabilities that R is $i > 0$ organized groups ahead, but $mi - j$ for $j \in [2, m]$ independents vote for L. Since $\chi^n > 0$, this inequality is impossible when v_L is close to v_I. It follows that, for any n, there is a ϑ_n such that if $\vartheta < \vartheta_n$ the proposition is satisfied. Note that ϑ_n is specific to a given equilibrium and a specific level of n: for an argument that holds for all possible equilibria in a large economy, we need a uniform bound over equilibria and n. To this goal, note:

$$\begin{aligned}
\chi^n &= \frac{1}{2}\sum_{i=1}^{\infty} S_i\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\cdot\sum_{j=2}^{m}\frac{(\sigma_n\alpha_I n)^{(mi-j)}e^{-\sigma_n\alpha_I n}}{(mi-j)!} \\
&> \frac{1}{2}S_1\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\cdot e^{-\sigma_n\alpha_I n}
\end{aligned}$$

In addition, note that condition (43) must hold in every equilibrium because independent voters mix in equilibrium. This implies that there exists a finite uniform upper bound $\overline{M}$ such that, for every n and every equilibrium $\sigma_n, \beta_R^n, \beta_L^n, T_j^n \leq \overline{M}$ for $j = I, L, R$ (details are presented in the online appendix). In the online appendix, we also establish a uniform lower bound on T_R^n. Specifically, we establish that there exist $\vartheta_1 > 0$, $N_1 < \infty$, and $\underline{M} > 0$ such that, for every $\vartheta \in [0, \vartheta_1]$, every $n \geq N_1$, and every equilibrium in Case 1, $T_R^n \geq \underline{M}$. From these two observations, it follows that $S_1\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right) \geq \min_{x\in[\underline{M},\overline{M}], y\in[0,\overline{M}]} S_1(x, y) > \underline{S}$ for some $\underline{S} > 0$. For any equilibrium and n, we have:

$$\chi^n \geq \frac{1}{2}S_1\left(\frac{\alpha_R\beta_R^n n}{m}, \frac{\alpha_L\beta_L^n n}{m}\right)\cdot e^{-\sigma_n\alpha_I n} > \frac{1}{2}\underline{S}\cdot e^{-\overline{M}} > 0,$$

On the other hand, $\frac{c}{v_L} - \frac{c}{v_I} = \frac{c\vartheta}{v_L(v_L+\vartheta)}$. Choose $\vartheta^* \leq \vartheta_1$ such that $\frac{c\vartheta^*}{v_L(v_L+\vartheta^*)} < \frac{1}{2}\underline{S}e^{-\overline{M}}$. Then, for every $\vartheta < \vartheta^*$, inequality (44) is impossible and the result is proven. Throughout this argument, the ordering between the two organized-group valuations is maintained and their values are fixed; the constants in the uniform bounds may therefore depend on these values.

Case 2: In Case 2, $\beta_R^n = 1$, $\beta_L^n \geq \frac{\alpha_R}{\alpha_L}$, and therefore $T_R^n = \frac{\alpha_R n}{m}$, $T_L^n \geq \frac{\alpha_R n}{m}$. In particular, $T_R^n \to \infty$. Equation (43) implies $\mathcal{P}_I(\sigma_n, \beta) \leq \sup_{i\in\mathbb{Z}} S_i(T_R^n, T_L^n)$. We have: $\sup_{i\in\mathbb{Z}} S_i(T_R^n, T_L^n) \leq \sup_{k\geq 0} \Pr\{\text{Pois}(T_R^n) = k\}$. The right-hand side converges to zero because $T_R^n \to \infty$. Hence, $\mathcal{P}_I(\sigma_n, \beta) \longrightarrow 0$. This contradicts the independent-voter indifference condition $\mathcal{P}_I(\sigma_n, \beta) = \frac{c}{v_I} > 0$. Thus Case 2 is impossible for all sufficiently large n. Combining this conclusion with the exclusion of Case 1 proves the proposition. ■

I.6 Proof of Proposition 5

If the voters use independent strategies (i.e. $m = 1$), turnout is distributed according to a Poisson with mean μ, and thus variance $\sigma_T^2 = \mu$, so $\sigma_T/\mu_T = \mu_T^{-1/2} \to 0$ as $\mu_T \to \infty$. Similarly, as noted in the text, we must have $\sigma_{MV}^2 = \mu_R + \mu_L$, so $\sigma_{MV}/\mu = 1/\sqrt{\mu} \to 0$ 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 re-profiles the leakage parameter κ across samples and specifications and shows that the central model comparison does not depend on its value. 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, on a single 90/10 split and across ten random splits. Throughout, the main estimation sample consists of the 1,703 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 $n < 2{,}000$, medium jurisdictions $2{,}000 \leq n < 8{,}500$, and large jurisdictions $n \geq 8{,}500$. The numerical procedures used here and throughout the paper—the likelihood implementation, the profiling algorithm, the county-block bootstrap, and the common scoring of the alternative models—are documented in detail in the Online Appendix (Appendix III.1). 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. The estimates are computed with the same profiled leakage parameter and county-block bootstrap as the full-sample model. As discussed in Section 5, the coefficients are stable across strata: the signs and magnitudes of the vote-ratio, turnout, and group-share loadings match the full-sample patterns, with the electorate-size elasticity of turnout between -0.35 and -0.38 across strata.

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 vote-ratio and turnout equations of the grouped model. Standard errors come from the county-block bootstrap described in the Online Appendix. The estimates confirm the patterns discussed in the main text: turnout declines steeply with electorate size, and the Baptist population share shifts the vote ratio toward the "against" side.

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 from the 1950s through the 1980s, 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 ($N = 501$), in 1970–1989 ($N = 569$), and in 1990 or later ($N = 633$), and re-estimate the complete grouped-vote model—all three equations—separately in each period. Each period is treated as a self-contained structural specification: the leakage parameter is profiled anew within the period, the coefficients are re-estimated at that period's $\hat{\kappa}$, and county-block bootstrap standard errors hold the period-specific $\hat{\kappa}$ fixed within each resample, paralleling the baseline inference. Table 11 reports the estimates and standard errors; the profiled leakage values are 0.059, 0.097, and 0.104 in the three periods, and the grouped model improves on the $m = 1$ benchmark by 385,565, 232,709, and 245,910 log-likelihood points, respectively.

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. The extended specification adds seven decade indicators (the pre-1950 period is omitted), an indicator for elections held concurrently with a larger election, standardized distance to the nearest college, and the standardized shares of adults aged 18–24 and 65 or older. The vote-ratio and turnout equations include all of these variables; the group-share equation adds the decade, distance, and age variables to its baseline electorate-size, city, Baptist-share, and congregation-size terms.

We estimate the extended model on two samples. The full extended sample preserves the electorate-size distribution that is central to the structural estimates. The first-referendum sample keeps only the first referendum observed in each locality ($N = 1{,}092$), and so verifies that the results are not driven by places that vote repeatedly. For each sample, the leakage parameter is profiled jointly with the full coefficient vector. Table 12 reports every coefficient in every equation, with county-block bootstrap standard errors for both the $m = 1$ and grouped specifications, and

Table 13 compares the resulting fits. The profiled leakage values are 0.099 in the full extended sample and 0.099 in the first-referendum sample, and the grouped model improves on the $m = 1$ benchmark by 859,546 and 413,006 log-likelihood points, respectively.

II.5 Re-profiling the compliance parameter

In this section, we show that the estimation results are robust to the choice of the compliance parameter κ, which is estimated by profiling the likelihood over a grid following the algorithm documented in Appendix III.1.3. Figure 9 plots the full-sample profile log likelihood. At each of the 100 grid points in $[0.01, 0.2]$, the complete grouped-vote coefficient vector is re-estimated, so every point on the curve is the best fit the grouped model can attain at that value of κ. The profile peaks at $\hat{\kappa} = 0.069$ with a log likelihood of $-36{,}814$ and falls to $-49{,}782$ at its least favorable grid point. The comparison with the $m = 1$ benchmark is therefore insensitive to κ in the strongest sense: the benchmark attains a log likelihood of $-1{,}023{,}634$, so even at the worst point of the grid the grouped model retains an advantage of more than 970,000 log-likelihood points. The entire profile curve lies far above the benchmark, not merely its peak.

Because the robustness exercises change the sample or the design matrix, κ is re-profiled by the same algorithm for each electorate-size stratum, each time period, each alternative eligible-voter floor, the CCM replication sample, and both extended-covariate samples. Table 14 collects the resulting point estimates; the re-profiled values range from 0.059 to 0.142 across specifications.

Table 15 reports the likelihood comparisons across these specifications, together with a sensitivity check on the compliance parameter: additional rows hold κ fixed at values well below and above the profiled optimum, re-estimating all coefficients at each value. In every specification the grouped model improves on the $m = 1$ model by at least 410,000 log-likelihood units, and by more than 900,000 in every full-sample specification.

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 group rule-utilitarian model can reproduce the dispersion of the margin of victory, its implied *turnout* distribution is qualitatively wrong. Figure 10 plots the model's simulated turnout density against the data: the model is sharply bimodal, with a spike near 7 percent turnout and a second mode near 30 percent, while the data display a single, smooth, right-skewed distribution. The bimodality follows directly from the model's participation side. Its two benefit indices depend almost entirely on the city-charter indicator—at the estimates, city status multiplies the two benefit parameters by factors of roughly 17 and 12—and every other

participation covariate is a binary indicator with a small loading, so fitted turnout concentrates on essentially two values, one for city and one for non-city elections. The support-share shock cannot fill the gap between them: a movement in μ reallocates votes between the two sides far more than it changes their sum, so it spreads the margin but adds almost no turnout dispersion. Within each city status the model predicts nearly the same turnout for every election, while observed turnout varies continuously from a few percent to more than eighty percent.

Second, Figure 11 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 is typically centered near 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. We randomly hold out 10 percent of the elections (170 of 1,703), re-estimate both the $m = 1$ benchmark and the grouped model from scratch on the remaining 90 percent (holding κ fixed at its full-sample profiled value), and score the fitted models on the held-out elections, which play no role in the estimation.

Table 16 reports the election-level prediction fit. On held-out turnout the two models perform comparably in terms of correlation (0.655 for the grouped model against 0.657 for $m = 1$), and the grouped model attains lower turnout prediction errors (mean absolute error 0.112 against 0.116). Figure 12 compares fitted and realized turnout election by election, in the estimation sample (left panel) and on the held-out elections (right panel): the quality of the fit is comparable in and out of sample, with no visible deterioration on the holdout.

The sharp out-of-sample distinction between the models concerns the margin of victory. Table 17 reports the dispersion and tail probabilities of the margin of victory on the held-out elections, comparing the empirical distribution with the distributions simulated from each model at the training-slice estimates. The grouped model reproduces the out-of-sample margin distribution

remarkably well: the standard deviation of the margin is 0.098 in the held-out data and 0.091 in the grouped model, against 0.034 under $m = 1$; margins of victory larger than 10 percent occur with probability 21.8% in the held-out data and 17.7% in the grouped model, against 1.2% under $m = 1$. The same ordering holds at every threshold. Figure 13 makes the point 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 virtually all mass in a narrow spike around a zero margin, while the grouped model tracks both the spread and the tails of the empirical distribution. The failure of the pivotal-voter model in the margin dimension is thus not repaired by estimation flexibility: it persists out of sample, and introducing organized groups repairs it.

The results above use a single fixed-seed random split. To verify that they are not an artifact of the particular split, we repeat the complete exercise—re-estimating both models on each training slice and scoring the corresponding holdout—for ten deterministic random seeds. The same split is used for the $m = 1$ and grouped models at each seed, and no holdout observation enters coefficient estimation at any point. Table 18 reports the mean and range of each holdout fit statistic across the ten splits, and Table 19 reports the cross-split distribution of the holdout margin-of-victory tails. The ordering of the models is unchanged in every split: the grouped model attains a higher mean holdout turnout correlation (0.657 against 0.643) and lower turnout errors (mean absolute error 0.116 against 0.121) than the $m = 1$ benchmark, its mean holdout margin-ratio correlation is comparable (0.178 against 0.181), and the tail ordering is completely stable—across the ten splits, $\Pr(|M| > 0.10)$ on the holdout ranges over $[0.14, 0.19]$ for the grouped model and $[0.006, 0.017]$ for $m = 1$, against empirical values in $[0.17, 0.22]$.

II.8 Tables

Table 9: Grouped-Vote Estimates by Electorate-Size Stratum

	Small	Medium	Large
Panel A: Vote-ratio equation, $\log r$			
Intercept	−0.0936	−0.0864	−0.1313***
		(0.0920)	(0.0406)
$\log(n)$	0.0141	0.0117	0.0151***
		(0.0107)	(0.0053)
Baptist share	−0.1801	−0.0510	−0.2434**
		(0.1622)	(0.1108)
City	0.0969	0.1066**	0.1546***
		(0.0439)	(0.0455)
Weekend	−0.0697	−0.0334	0.0062
		(0.0408)	(0.0383)
Alcohol (off-prem)	−0.1021	−0.1673***	−0.1164**
		(0.0451)	(0.0494)
Alcohol (on/off)	0.1059	0.0628	−0.0505
		(0.0546)	(0.0569)
Panel B: Turnout equation, $\text{logit}(t)$			
Intercept	0.9429	0.9084***	0.7010***
		(0.1693)	(0.0793)
$\log(n)$	−0.3769	−0.3555***	−0.3551***
		(0.0181)	(0.0094)
Baptist share	1.9999	2.2332***	2.5862***
		(0.2971)	(0.2694)
City	0.6681	0.6324***	1.0576***
		(0.0775)	(0.1391)
Weekend	−0.2140	−0.1804***	−0.2533***
		(0.0588)	(0.0290)
Alcohol (off-prem)	0.0390	−0.1548*	−0.0847
		(0.0803)	(0.0615)
Alcohol (on/off)	0.1482	−0.0309	0.0884
		(0.0798)	(0.0815)
Panel C: Group-share equation, $\log(m/n)$			
Intercept	1.8948	1.9072***	1.9759***
		(0.0191)	(0.0334)
$\log(n)$	−0.8396	−0.8061***	−0.7964***
		(0.0048)	(0.0037)
$[\log(n)]^2$	0.0071	0.0070***	0.0075***
		(0.0003)	(0.0003)
City	−0.2826	−0.2686***	−0.2752***
		(0.0086)	(0.0137)
Baptist share	−0.0450	−0.0277*	−0.0198
		(0.0152)	(0.0391)
SBC congregation size (std.)	0.0684	0.0982***	0.1132***
		(0.0100)	(0.0247)
N	575	592	536
Log likelihood	−6941.8	−11301.8	−17664.9
Mean $\hat{m}$	19.54	51.29	118.23
Median $\hat{m}$	19.00	52.00	96.00

Notes: Maximum-likelihood estimates for the grouped-vote model within each electorate-size stratum. The leakage parameter is fixed at $\kappa = 0.069$. The group equation uses standardized Southern Baptist average congregation size in addition to the baseline covariates shown in Panel C. County-block bootstrap standard errors (200 draws) are in parentheses. *, **, and *** denote significance at the 10, 5, and 1 percent levels. Full-sample estimates are reported in Table 7.

Table 10: Main Benchmark Estimates: $m = 1$

	Full	Small	Medium	Large
Panel A: Vote-ratio equation, $\log r$				
Intercept	−0.1346	0.0540	0.0982	−0.1952
	(0.1696)	(0.2142)	(0.3340)	(0.3618)
$\log(n)$	0.0129	−0.0022	−0.0057	0.0148
	(0.0212)	(0.0259)	(0.0385)	(0.0392)
Baptist share	−0.2222	−0.1983	−0.4040***	−0.1111
	(0.1696)	(0.1665)	(0.1563)	(0.2443)
City	0.2126***	0.1431***	0.1754***	0.2343***
	(0.0328)	(0.0492)	(0.0368)	(0.0358)
Weekend	−0.0456	−0.1249***	−0.0626**	−0.0345
	(0.0359)	(0.0297)	(0.0294)	(0.0468)
Alcohol (off-prem)	−0.0808**	−0.1111***	−0.0855***	−0.0561
	(0.0338)	(0.0312)	(0.0282)	(0.0459)
Alcohol (on/off)	−0.0376	0.0741*	−0.0116	−0.0593
	(0.0312)	(0.0449)	(0.0359)	(0.0435)
Panel B: Turnout equation, $\text{logit}(t)$				
Intercept	0.5525	0.9295*	−0.6631	1.5390*
	(0.5804)	(0.5468)	(1.0808)	(0.8844)
$\log(n)$	−0.3125***	−0.3246***	−0.1758	−0.3998***
	(0.0486)	(0.0665)	(0.1276)	(0.0732)
Baptist share	2.3063***	1.2325***	2.2172***	2.2484**
	(0.7479)	(0.3695)	(0.3977)	(1.0220)
City	0.9732***	0.2994*	0.6352***	1.1106***
	(0.1315)	(0.1622)	(0.1054)	(0.1749)
Weekend	−0.3942***	−0.0731	−0.1703	−0.5228***
	(0.1164)	(0.0855)	(0.1038)	(0.1353)
Alcohol (off-prem)	−0.1428	0.2488***	0.0357	−0.2139
	(0.1032)	(0.0797)	(0.1153)	(0.1572)
Alcohol (on/off)	0.1247	−0.0447	−0.2727***	0.2895*
	(0.1406)	(0.1115)	(0.0962)	(0.1610)
N	1,703	575	592	536
Log likelihood	−1023634.2	−26256.5	−161197.3	−802178.5

Notes: Maximum-likelihood estimates for the independent-Poisson benchmark. The main sample is the previously dry Texas sample with non-missing baseline covariates, turnout no greater than one, and at least 100 eligible voters. Small, medium, and large strata contain elections with $n < 2{,}000$, $2{,}000 \leq n < 8{,}500$, and $n \geq 8{,}500$ eligible voters. County-block bootstrap standard errors (200 draws) are in parentheses. *, **, and *** denote significance at the 10, 5, and 1 percent levels.

Table 11: Grouped-Vote Estimates Re-estimated by Time Period

	Pre-1970	1970–1989	1990+
Panel A: Vote-ratio equation, $\log r$			
Intercept	-0.4042^{***}	-0.0552	0.0841^{**}
	(0.0103)	(0.0503)	(0.0421)
$\log(n)$	0.0328^{***}	0.0161^{***}	0.0139^{***}
	(0.0049)	(0.0052)	(0.0043)
Baptist share	-0.4012^{***}	-0.8567^{***}	-0.2997^{***}
	(0.0984)	(0.1010)	(0.0900)
City	0.0938^{*}	0.1125^{***}	0.0469
	(0.0512)	(0.0365)	(0.0302)
Weekend	-0.0123	0.0427	0.0298
	(0.0350)	(0.0393)	(0.0345)
Alcohol (off-premise)	0.0269	-0.0135	-0.0569
	(0.0438)	(0.0380)	(0.0411)
Alcohol (on/off-premise)	0.0954^{**}	0.0164	-0.0195
	(0.0433)	(0.0440)	(0.0470)
Panel B: Turnout equation, $\text{logit}(t)$			
Intercept	0.7033^{***}	0.2622^{***}	0.9360^{***}
	(0.0916)	(0.0682)	(0.0566)
$\log(n)$	-0.3887^{***}	-0.3762^{***}	-0.2749^{***}
	(0.0068)	(0.0090)	(0.0050)
Baptist share	4.1683^{***}	2.7802^{***}	0.4574^{***}
	(0.2747)	(0.3076)	(0.1294)
City	0.8545^{***}	1.2387^{***}	0.5828^{***}
	(0.0892)	(0.1133)	(0.0516)
Weekend	0.0274	-0.0196	-0.6385^{***}
	(0.0563)	(0.0534)	(0.0079)
Alcohol (off-premise)	-0.3428^{***}	0.0611	-0.1046
	(0.0353)	(0.0554)	(0.0674)
Alcohol (on/off-premise)	0.2170^{***}	0.2975^{***}	-0.1216^{*}
	(0.0813)	(0.0844)	(0.0701)
Panel C: Group-share equation, $\log(m/n)$			
Intercept	1.8722^{***}	1.8603^{***}	1.6616^{***}
	(0.0203)	(0.0167)	(0.0128)
$\log(n)$	-0.8049^{***}	-0.8002^{***}	-0.8019^{***}
	(0.0027)	(0.0022)	(0.0020)
$[\log(n)]^2$	0.0072^{***}	0.0080^{***}	0.0137^{***}
	(0.0002)	(0.0001)	(0.0002)
City	-0.2618^{***}	-0.2625^{***}	-0.2716^{***}
	(0.0186)	(0.0097)	(0.0065)
Baptist share	-0.1079^{***}	0.0111	-0.2308^{***}
	(0.0393)	(0.0260)	(0.0336)
SBC congregation size (std.)	0.1687^{***}	0.1589^{***}	0.0861^{***}
	(0.0153)	(0.0167)	(0.0152)
Profiled $\hat{\kappa}$	0.059	0.097	0.104
N	501	569	633
LL grouped	-12508.6	-10724.4	-10568.2
LL $m = 1$	-398073.2	-243433.7	-256478.4
ΔLL	385564.6	232709.4	245910.1
Mean $\hat{m}$	70.60	62.69	64.78
	(5.80)	(4.41)	(5.57)
Median $\hat{m}$	57.00	53.00	42.00
	(3.13)	(3.42)	(2.91)

Notes: Each column re-estimates the complete grouped-vote model on the elections of the indicated period. The leakage parameter is held fixed at the period-specific profiled value $\hat{\kappa}$ reported in the bottom panel, both for the point estimates and within every county-block bootstrap replicate. Bootstrap standard errors appear in parentheses below each estimate. *, **, and *** denote significance at the 10, 5, and 1 percent levels.

Table 12: Extended-Covariate Estimates: All Coefficients

Covariate	Full $m = 1$	Full grouped	First-ref. $m = 1$	First-ref. grouped
Panel A: Vote-ratio equation, $\log r$				
	-0.4045^{**}	-0.0993^{***}	-0.3945^{***}	-0.1288^{***}
Intercept	(0.1652)	(0.0045)	(0.1137)	(0.0066)
	0.0144	-0.0011^{*}	0.0186	0.0013
$\log(n)$	(0.0206)	(0.0006)	(0.0113)	(0.0010)
	0.0629	-0.1713^{***}	-0.3524^{**}	-0.4105^{***}
Baptist share	(0.1915)	(0.0176)	(0.1628)	(0.0224)
	0.0345	0.0027	0.0448	-0.0192^{*}
City	(0.0331)	(0.0077)	(0.0281)	(0.0110)
	0.0099	-0.1216^{***}	-0.0466^{*}	-0.0913^{***}
Weekend	(0.0275)	(0.0053)	(0.0280)	(0.0107)
	-0.0266	-0.0448^{***}	-0.0543	-0.0123
Alcohol (off-premise)	(0.0342)	(0.0095)	(0.0393)	(0.0134)
	-0.0238	0.0629^{***}	0.0258	0.0103
Alcohol (on/off-premise)	(0.0268)	(0.0140)	(0.0248)	(0.0220)
	0.0035	0.0283	0.0688	0.1710^{***}
1950s	(0.0460)	(0.0270)	(0.0669)	(0.0446)
	0.1085^{**}	0.2441^{***}	0.2355^{***}	0.2914^{***}
1960s	(0.0469)	(0.0187)	(0.0512)	(0.0254)
	0.1804^{***}	0.1866^{***}	0.2378^{***}	0.2307^{***}
1970s	(0.0432)	(0.0166)	(0.0405)	(0.0217)
	0.2138^{***}	0.2317^{***}	0.3078^{***}	0.3034^{***}
1980s	(0.0444)	(0.0153)	(0.0518)	(0.0254)
	0.3027^{***}	0.2201^{***}	0.4059^{***}	0.3921^{***}
1990s	(0.0527)	(0.0328)	(0.0680)	(0.0458)
	0.3832^{***}	0.3710^{***}	0.5212^{***}	0.4943^{***}
2000s	(0.0386)	(0.0202)	(0.0550)	(0.0218)
	0.5270^{***}	0.4272^{***}	0.5874^{***}	0.5760^{***}
2010s	(0.0476)	(0.0206)	(0.0599)	(0.0212)
	-0.0600^{*}	-0.0753^{***}	-0.0636	-0.1437^{***}
Concurrent election	(0.0359)	(0.0121)	(0.0508)	(0.0143)
	-0.0018	0.0248^{***}	0.0158	0.0283^{***}
Distance to college (std.)	(0.0157)	(0.0063)	(0.0157)	(0.0092)
	0.0046	0.0085^{*}	0.0009	0.0143^{*}
Share age 18–24 (std.)	(0.0167)	(0.0050)	(0.0143)	(0.0085)
	-0.0305	-0.0256^{***}	-0.0103	-0.0305^{***}
Share age 65+ (std.)	(0.0192)	(0.0052)	(0.0182)	(0.0077)
Panel B: Turnout equation, $\text{logit}(t)$				
	0.6730	0.7009^{***}	0.5021	0.7095^{***}
Intercept	(0.6069)	(0.0124)	(0.5717)	(0.0183)
	-0.3021^{***}	-0.3315^{***}	-0.3276^{***}	-0.3729^{***}
$\log(n)$	(0.0509)	(0.0015)	(0.0509)	(0.0030)
	1.8747^{**}	2.4197^{***}	1.1350	1.8061^{***}
Baptist share	(0.8556)	(0.0466)	(0.7449)	(0.0676)
	1.2092^{***}	1.1166^{***}	1.4403^{***}	1.1945^{***}
City	(0.1805)	(0.0323)	(0.1861)	(0.0623)
	-0.0976	-0.0145	-0.1204	-0.0325^{***}
Weekend	(0.1314)	(0.0114)	(0.1864)	(0.0100)
	0.0075	-0.0904^{***}	0.0264	-0.0521^{***}
Alcohol (off-premise)	(0.1171)	(0.0113)	(0.1495)	(0.0173)
	0.2360^{**}	0.1388^{***}	0.2740^{*}	0.1069^{***}
Alcohol (on/off-premise)	(0.1098)	(0.0170)	(0.1520)	(0.0318)
	-0.0805	-0.1676^{***}	-0.9016^{***}	-0.5125^{***}
1950s	(0.2185)	(0.0282)	(0.2808)	(0.0398)
	-0.9445^{***}	-0.7790^{***}	-0.5718^{**}	-0.2486^{***}
1960s	(0.2662)	(0.0096)	(0.2483)	(0.0240)
	-0.6123^{***}	-0.5172^{***}	0.0289	-0.0855^{***}
1970s	(0.1563)	(0.0084)	(0.2028)	(0.0230)
	-0.9202^{***}	-0.6525^{***}	-0.2286	-0.1076^{***}
1980s	(0.2567)	(0.0070)	(0.3289)	(0.0292)
	-1.0446^{***}	-0.6715^{***}	-0.6211^{*}	-0.4017^{***}
1990s	(0.2753)	(0.0254)	(0.3321)	(0.0391)
	-0.6785^{***}	-0.5392^{***}	-0.4117	-0.1779^{***}
2000s	(0.2068)	(0.0089)	(0.2856)	(0.0192)

continued on next page

Table 12, continued

Covariate	Full $m = 1$	Full grouped	First-ref. $m = 1$	First-ref. grouped
	-1.0279^{***}	-1.0742^{***}	-0.4679	-0.5838^{***}
2010s	(0.2421)	(0.0022)	(0.3317)	(0.0097)
	0.6472^{***}	0.6201^{***}	0.4676^{**}	0.5433^{***}
Concurrent election	(0.1925)	(0.0329)	(0.2070)	(0.0391)
	-0.0664	0.0229^{***}	-0.0695	0.0148
Distance to college (std.)	(0.0738)	(0.0087)	(0.0940)	(0.0111)
	0.0176	0.0187^{*}	0.0355	0.0334^{**}
Share age 18–24 (std.)	(0.1128)	(0.0106)	(0.0992)	(0.0150)
	0.1565	0.1092^{***}	0.1602	0.1374^{***}
Share age 65+ (std.)	(0.0982)	(0.0102)	(0.1013)	(0.0156)
Panel C: Group-share equation, $\log(m/n)$ *(grouped model only)*				
		1.8686^{***}		1.8794^{***}
Intercept	–	(0.0061)	–	(0.0065)
		-0.8017^{***}		-0.8113^{***}
$\log(n)$	–	(0.0007)	–	(0.0008)
		0.0083^{***}		0.0077^{***}
$[\log(n)]^2$	–	(0.0001)	–	(0.0001)
		-0.2819^{***}		-0.2661^{***}
City	–	(0.0043)	–	(0.0043)
		-0.0474^{***}		-0.0468^{**}
Baptist share	–	(0.0134)	–	(0.0189)
		0.0518^{***}		0.0141
1950s	–	(0.0171)	–	(0.0208)
		-0.0837^{***}		0.0405
1960s	–	(0.0195)	–	(0.0256)
		0.0006		0.0913^{***}
1970s	–	(0.0093)	–	(0.0116)
		-0.0442^{***}		0.0371^{***}
1980s	–	(0.0158)	–	(0.0126)
		0.0649^{**}		-0.1352^{***}
1990s	–	(0.0270)	–	(0.0145)
		0.2147^{***}		0.1120^{***}
2000s	–	(0.0215)	–	(0.0149)
		0.1348^{***}		0.0192
2010s	–	(0.0184)	–	(0.0159)
		-0.0248^{***}		-0.0287^{***}
Distance to college (std.)	–	(0.0068)	–	(0.0061)
		-0.0196^{**}		0.0335^{***}
Share age 18–24 (std.)	–	(0.0080)	–	(0.0066)
		-0.0095^{**}		-0.0220^{***}
Share age 65+ (std.)	–	(0.0045)	–	(0.0052)
		0.1311^{***}		0.1633^{***}
SBC congregation size (std.)	–	(0.0057)	–	(0.0096)

Notes: Cells report point estimates with county-block bootstrap standard errors in parentheses beneath. *, **, and *** denote significance at the 10, 5, and 1 percent levels. Columns pair the $m = 1$ and grouped specifications on the full extended sample and on the first-referendum sample. Grouped columns hold the leakage parameter fixed at the specification-specific profiled value $\hat{\kappa}$; the $m = 1$ model has no leakage parameter, and the group-share equation exists only in the grouped model. Weekend, alcohol-type, and concurrent-election indicators enter the vote-ratio and turnout equations only, and the omitted time category is elections held before 1950.

Table 13: Extended-Covariate Fits

Sample	N	Profiled $\hat{\kappa}$	LL($m = 1$)	LL(grouped)	ΔLL	Mean $\hat{m}$	Bootstrap success ($m = 1$/G)
Full sample	1,703	0.083	−895960.4	−34789.7	861170.7	67.12	200/200 / 94/200
First referendum	1,092	0.091	−435138.5	−20865.4	414273.1	59.43	200/200 / 131/200

Notes: The grouped model in each row holds the compliance parameter fixed at that specification's profiled value $\hat{\kappa}$, both for estimation and within the county-block bootstrap. The final column reports the number of bootstrap replicates used over the target, for the $m = 1$ and grouped models respectively.

Table 14: Profile Estimates of the Leakage Parameter Across Specifications

Specification	N	$\hat{\kappa}$	Log likelihood
Main-sample small stratum	575	0.1226	−6677.1
Main-sample medium stratum	592	0.1307	−11070.7
Main-sample large stratum	536	0.0637	−17114.1
Pre-1970	501	0.0586	−12508.6
1970–1989	569	0.0973	−10724.4
1990+	633	0.1040	−10568.2
Eligible-voter floor: $n \geq 50$	1,710	0.0698	−36869.6
Eligible-voter floor: $n \geq 200$	1,642	0.0682	−36161.4
Eligible-voter floor: $n \geq 500$	1,491	0.0643	−34067.2
Extended covariates: full sample	1,703	0.0829	−34789.7
Extended covariates: first referendum	1,092	0.0906	−20865.4
CCM sample: $n \leq 900$	141	0.1417	−1402.7

Notes: Each row separately profiles the grouped-vote likelihood over κ, re-estimating every coefficient at each grid point. The reported point estimate is the best of a multi-start joint maximum-likelihood stage, repeated ascending and descending sweeps over the grid, and a final joint polish.

Table 15: Robustness Checks

Specification	N	LL($m = 1$)	LL(grouped)	ΔLL	ΔLL/N	Mean $\hat{m}$
Panel A: Alternative samples and electorate-size thresholds						
Texas, $n \geq 50$	1,710	−1023750.0	−36869.6	986880.4	577.123	59.17
Texas, $n \geq 200$	1,642	−1022494.7	−36161.4	986333.3	600.690	60.80
Texas, $n \geq 500$	1,491	−1017713.8	−34067.2	983646.5	659.723	65.53
Full sample, extended covariates	1,703	−895960.4	−34789.7	861170.7	505.679	67.12
First referendum, extended covariates	1,092	−435138.5	−20865.4	414273.1	379.371	59.43
CCM, $n \leq 900$	141	−3304.8	−1402.7	1902.2	13.491	18.35
Panel B: Fixed-κ sensitivity, Texas baseline sample						
$\kappa = 0.020$	1,703	−1023634.2	−43147.7	980486.6	575.741	41.68
$\kappa = 0.050$	1,703	−1023634.2	−37447.5	986186.7	579.088	59.15
$\kappa = 0.069$	1,703	−1023634.2	−36813.1	986821.2	579.460	59.12
$\kappa = 0.150$	1,703	−1023634.2	−38292.6	985341.7	578.592	84.26

Notes: In Panel A, grouped-model log likelihoods and mean group sizes are evaluated at the specification-specific profile optima of Table 14; the $m = 1$ likelihood is estimated on the identical sample and covariate matrix. The main $n \geq 100$ sample is reported separately in Table 20. Panel B re-estimates the grouped model on the Texas baseline sample with the compliance parameter held fixed at the stated value: active-group members fail to execute independently with probability κ, and a balancing Poisson($\kappa m B \theta_s$) flow of individual votes arrives, so expected votes are unchanged. The CCM row uses elections with at most 900 eligible voters.

Table 16: Re-estimated Holdout Prediction Fit

Model	Sample	Elections	Corr. T	Corr. R/L	MAE T	MAE R/L
$m = 1$	Training (90%)	1,526	0.652	0.165	0.120	0.735
$m = 1$	Holdout (10%)	170	0.657	0.391	0.116	0.652
Grouped	Training (90%)	1,526	0.661	0.131	0.116	0.740
Grouped	Holdout (10%)	170	0.655	0.416	0.112	0.658

Notes: Models are re-estimated on a fixed-seed random 90% training slice and scored on the held-out 10% of elections. The grouped model fixes $\kappa = 0.069$. The margin ratio is the direct side ratio: V_R/V_L empirically and $\hat{\theta}_R/\hat{\theta}_L$ in prediction.

Table 17: Holdout Margin-of-Victory Distribution

Series	SD margin	$\Pr(\lvert M\rvert > 0.05)$	$\Pr(\lvert M\rvert > 0.10)$	$\Pr(\lvert M\rvert > 0.15)$	$\Pr(\lvert M\rvert > 0.20)$
Empirical	0.098	0.382	0.218	0.135	0.065
$m = 1$	0.034	0.134	0.012	0.001	0.000
Grouped	0.091	0.378	0.177	0.095	0.050

Notes: Margin-of-victory dispersion and tail probabilities on the held-out 10% of elections. Both models are re-estimated on the 90% training slice only; model rows simulate 250 elections per holdout observation at the training-slice estimates. The grouped model fixes $\kappa = 0.069$. M is the margin of victory $(V_R - V_L)/n$.

Table 18: Holdout Prediction Fit Across 10 Random Splits

Statistic	$m = 1$	Grouped
Corr. T	0.643 [0.594, 0.712]	0.657 [0.613, 0.734]
Corr. R/L	0.181 [0.041, 0.391]	0.178 [0.020, 0.416]
MAE T	0.121 [0.111, 0.133]	0.116 [0.105, 0.127]
MAE R/L	0.677 [0.522, 0.881]	0.680 [0.528, 0.882]
RMSE T	0.158 [0.142, 0.172]	0.151 [0.133, 0.164]
RMSE R/L	1.465 [0.937, 2.290]	1.458 [0.933, 2.281]

Notes: Each split re-estimates both models on a random 90% training slice and scores the held-out 10%. Cells report the mean across 10 independent splits with the [min, max] range in brackets. The grouped model fixes $\kappa = 0.069$ in every split. The margin ratio is V_R/V_L empirically and $\hat{\theta}_R/\hat{\theta}_L$ in prediction.

Table 19: Holdout Margin-of-Victory Distribution Across 10 Random Splits

Statistic	Empirical	$m = 1$	Grouped
SD margin	0.093 [0.079, 0.111]	0.034 [0.030, 0.037]	0.090 [0.078, 0.101]
$\Pr(\lvert M\rvert > 0.05)$	0.385 [0.341, 0.412]	0.120 [0.084, 0.151]	0.368 [0.331, 0.391]
$\Pr(\lvert M\rvert > 0.10)$	0.199 [0.171, 0.218]	0.011 [0.006, 0.017]	0.168 [0.137, 0.189]
$\Pr(\lvert M\rvert > 0.15)$	0.107 [0.071, 0.141]	0.001 [0.000, 0.002]	0.087 [0.066, 0.103]
$\Pr(\lvert M\rvert > 0.20)$	0.059 [0.035, 0.088]	0.000 [0.000, 0.000]	0.048 [0.034, 0.063]

Notes: Cells report the mean across 10 independent 90/10 splits with the [min, max] range in brackets. The Empirical row varies across splits because each split holds out a different random 10% of elections.

II.9 Figures

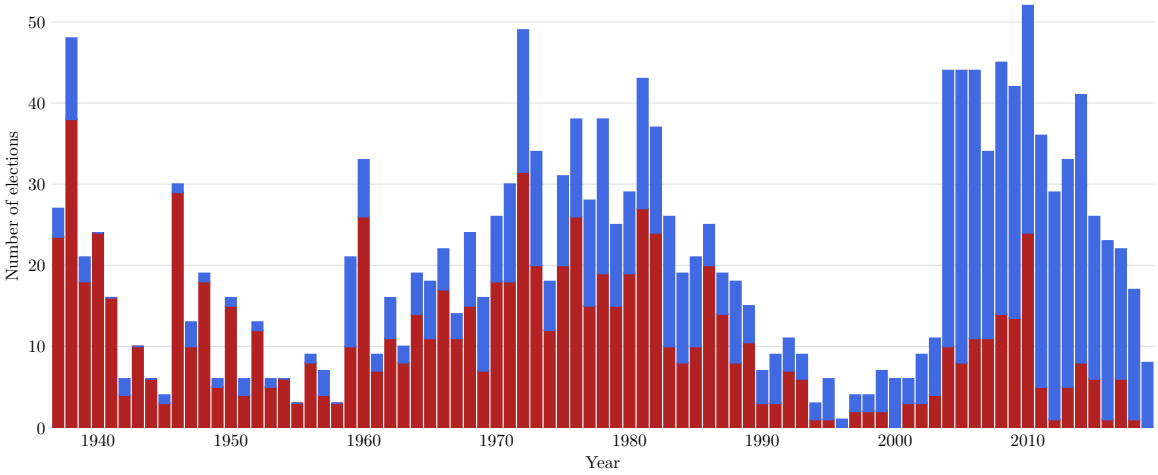


Figure 7: Liquor referenda in previously-dry jurisdictions over time. The red lower segment of each bar records elections in which "against" won (the jurisdiction stayed dry); the blue upper segment records elections in which "for" won.

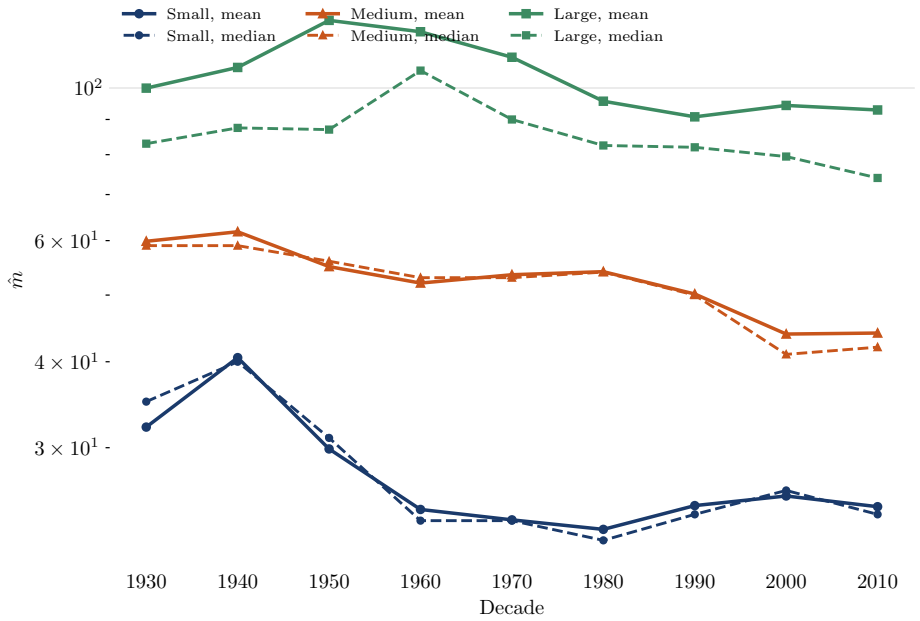


Figure 8: Fitted group size by decade and electorate-size stratum, from the main full-sample grouped model. The figure is descriptive; the period-specific structural estimates are reported in Table 11.

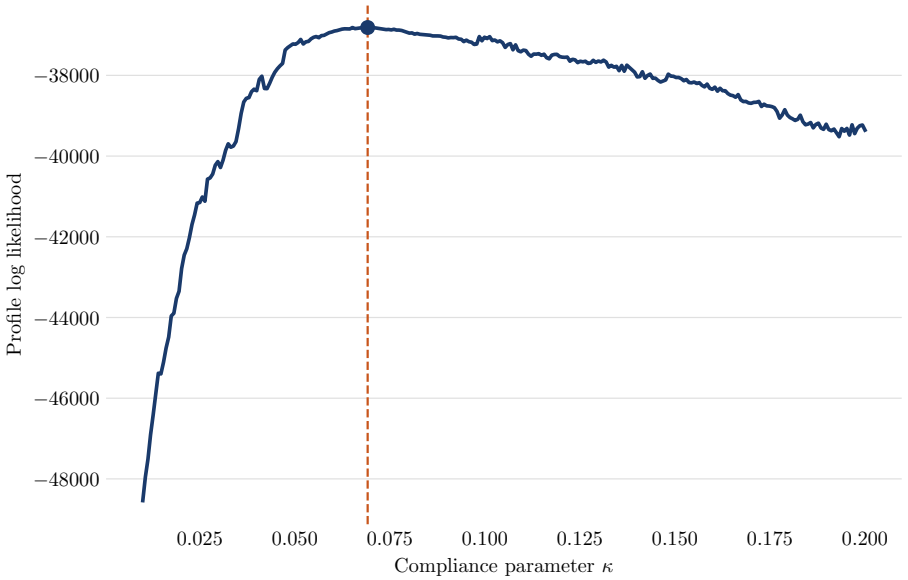


Figure 9: Profile log likelihood for κ in the full Texas sample. At every grid point the complete grouped-vote coefficient vector is re-estimated. The profile peaks at $\hat{\kappa} = 0.069$; for comparison, the $m = 1$ benchmark attains $-1{,}023{,}634$, more than 970,000 log-likelihood points below every point of the curve.

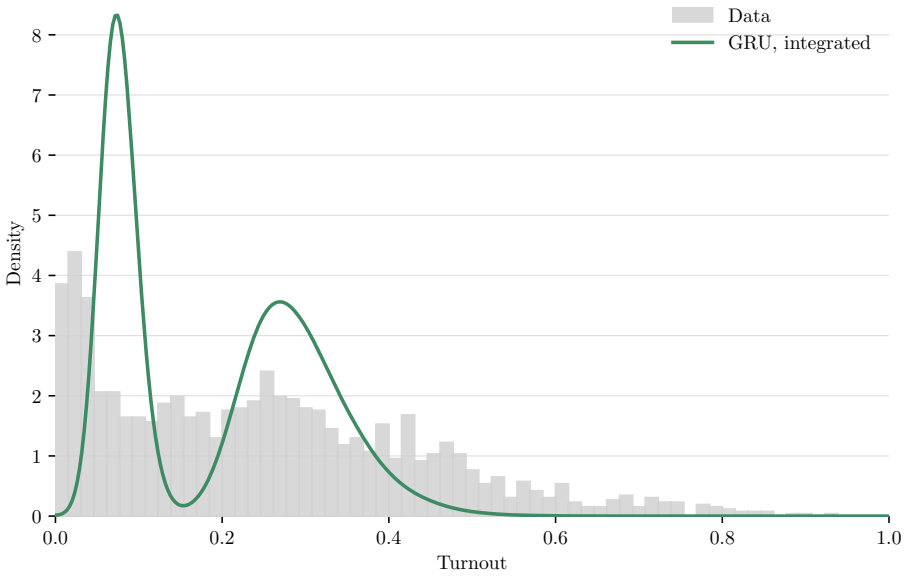


Figure 10: Simulated turnout distribution under the group rule-utilitarian model (green line) against the data (histogram).

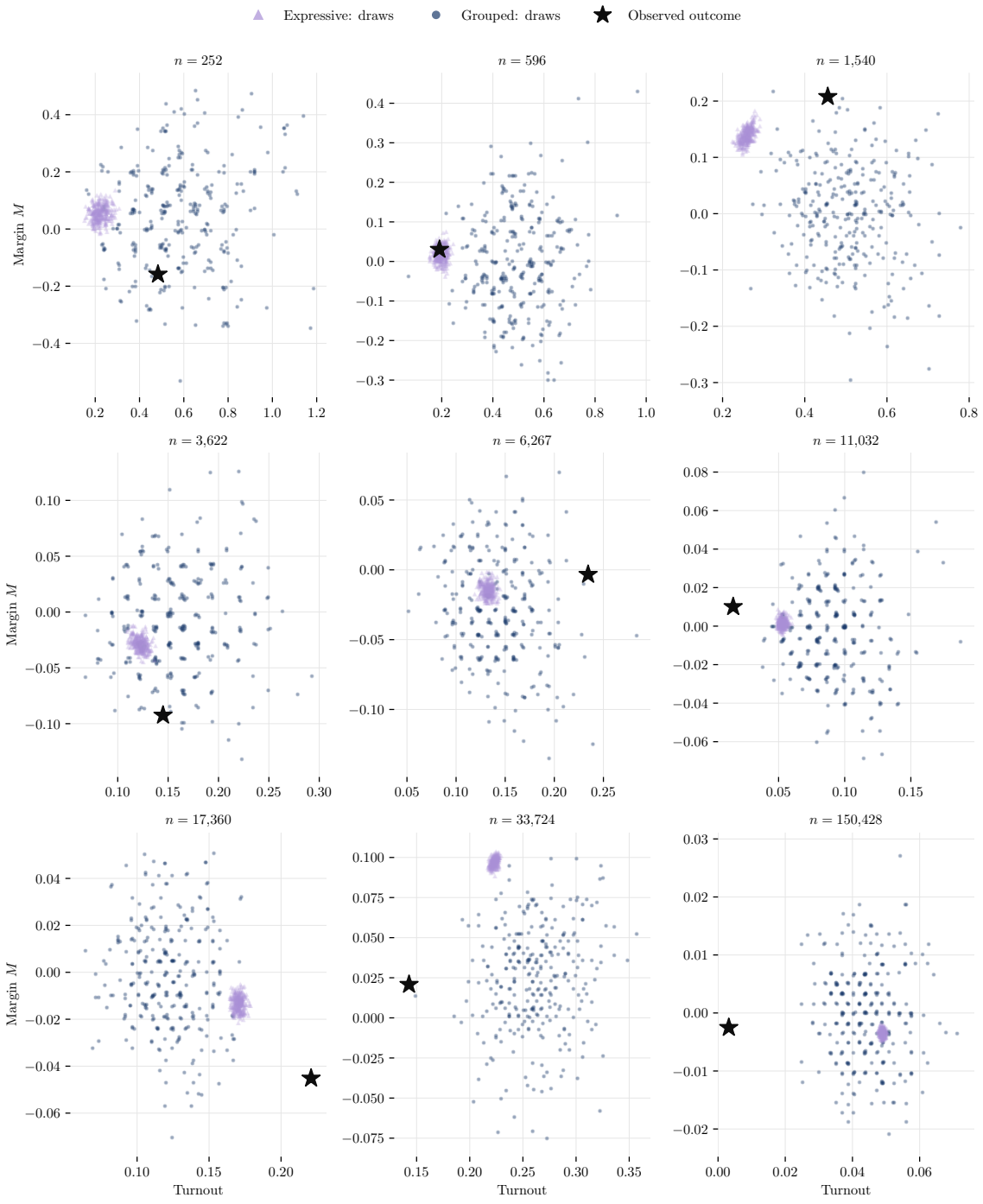


Figure 11: Election-level predictive distributions under the expressive-voting model (purple) and the grouped model (blue), for nine elections at fixed quantiles of the electorate size; the star marks the observed outcome. Each panel shows 300 simulated draws from each model.

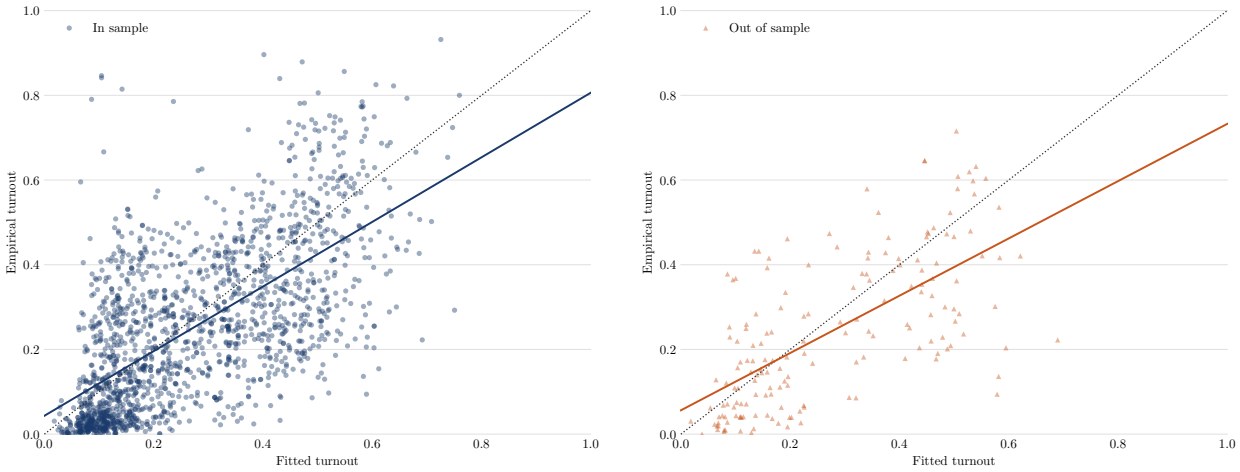


Figure 12: Fitted versus empirical turnout under the grouped model: full estimation sample (left) and held-out elections at the training-slice estimates (right). The dotted line is the 45-degree line.

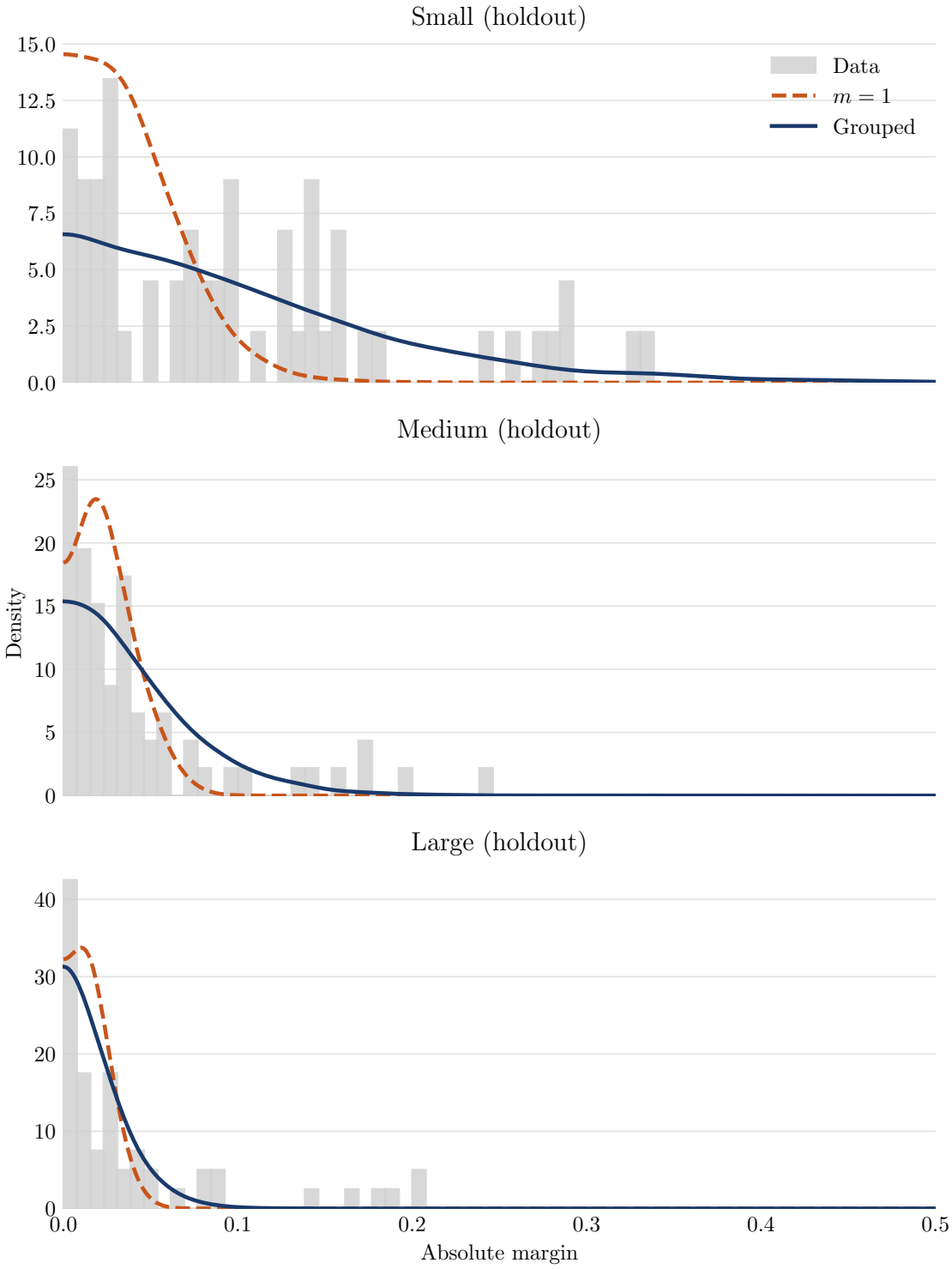


Figure 13: Absolute margin of victory out of sample, by electorate-size stratum. Panels are stacked by stratum (small, medium, and large jurisdictions); in each panel the empirical distribution of the absolute margin of victory on the held-out 10% of elections is compared with simulations from the $m = 1$ benchmark and the grouped-vote model, both re-estimated on the 90% training slice only (the out-of-sample exercise of 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 (2017): "Voter Turnout and Preference Aggregation," Mimeo, Cornell University and Yale University.

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.

Clarke, K. A. (2007): "A Simple Distribution-Free Test for Nonnested Model Selection," *Political Analysis*, 15, 347–363.

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, 408–204.

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 (1987): "Expressive Political Behavior: Foundations, Scope, and Implications," *British Journal of Political Science*, 4, 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 and Y. Watanabe (2013): "Inferring Strategic Voting," *American Economic Review*, 103, 624–662.

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

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.

Liu, D. C. and J. Nocedal (1989): "On the Limited Memory BFGS Method for Large Scale Optimization," *Mathematical Programming*, 45, 503–528.

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.

——— (2000): "Large Poisson games," *Journal of Economic Theory*, 94, 375–392.

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

Nelder, J. A. and R. Mead (1965): "A Simplex Method for Function Minimization," *The Computer Journal*, 7, 308–313.

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.

Skellam, J. G. (1946): "The Frequency Distribution of the Difference Between Two Poisson Variates Belonging to Different Populations," *Journal of the Royal Statistical Society*, 109, 296.

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.

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307–333.

III Appendix C: Online Appendix

III.1 Estimation and Implementation Details (Online Appendix)

This appendix documents the estimation machinery behind the results in the paper: the parameterization and likelihood (Section III.1.1), the numerical evaluation and optimization strategy (Section III.1.2), the algorithm that profiles the leakage parameter (Section III.1.3), the county-block bootstrap (Section III.1.4), and the simulation and common scoring of the alternative voting models (Section III.1.5).

III.1.1 Sample, parameterization, and likelihood

For election i, let n_i denote the number of eligible voters and $(V_{R,i}, V_{L,i})$ the integer vote totals for and against legalization. Rather than parameterizing the two side-specific turnout intensities $(\theta_{R,i}, \theta_{L,i})$ directly, we work with the vote ratio and total turnout,

$$r_i = \sqrt{\theta_{R,i}/\theta_{L,i}}, \qquad t_i = \theta_{R,i} + \theta_{L,i},$$

with index functions

$$\log r_i = X'_{r,i}\beta_r, \qquad \operatorname{logit}(t_i) = X'_{t,i}\beta_t.$$

Inverting the transformation gives

$$\theta_{R,i} = \frac{r_i^2 t_i}{1 + r_i^2}, \qquad \theta_{L,i} = \frac{t_i}{1 + r_i^2}.$$

This parameterization has two advantages. First, it keeps both intensities strictly positive and their sum strictly below one at every point the optimizer can visit, so no constraint handling is needed inside the likelihood. Second, it separates the two economically distinct margins: the vote-ratio equation governs which side is favored, and the turnout equation governs how many citizens participate. The same design matrix is used in both equations.

In the $m = 1$ benchmark, the vote totals are independent Poisson random variables,

$$V_{R,i} \sim \text{Poisson}(n_i\theta_{R,i}), \qquad V_{L,i} \sim \text{Poisson}(n_i\theta_{L,i}),$$

and the election-level log likelihood is the sum of the two Poisson log probabilities.

In the grouped model, latent group size is parameterized as a share of the electorate,

$$\log(m_i/n_i) = X'_{m,i}\beta_m,$$

which keeps group size meaningful across electorates that differ in size by three orders of magnitude. The continuous index is clamped to the feasible interval and deterministically rounded to an integer

$m_i \in \{1, \ldots, n_i\}$, and the number of potential groups is $B_i = \lfloor n_i/m_i \rfloor$. For side $s \in \{R, L\}$, observed votes are the sum of a block component and a residual individual component,

$$V_{s,i} = m_i Y_{s,i} + U_{s,i}, \qquad Y_{s,i} \sim \text{Poisson}(B_i \theta_{s,i}), \qquad U_{s,i} \sim \text{Poisson}(\kappa n_i \theta_{s,i}),$$

so the probability of a vote total is the finite convolution

$$\Pr(V_{s,i} = v) = \sum_{y=0}^{\lfloor v/m_i \rfloor} p(y; B_i \theta_{s,i})\, p(v - m_i y; \kappa n_i \theta_{s,i}),$$

where $p(\cdot; \lambda)$ denotes a Poisson probability mass function. Conditional on the covariates, the two sides are independent, so the election-level likelihood is the product of the two side probabilities. The leakage component serves a mechanical purpose—without it, any vote total that is not an exact multiple of m_i would receive probability zero—but its estimated size is disciplined by the data through the profiling procedure of Section III.1.3.

III.1.2 Numerical evaluation and optimization

The compound-Poisson sum is evaluated exactly in log space: factorials are computed with log-gamma functions, and the terms of the convolution are aggregated with log-sum-exp, which prevents underflow in large electorates where individual Poisson probabilities can be smaller than the smallest representable floating-point number. An invalid numerical evaluation receives a large finite objective penalty rather than an infinite or missing value, so the optimizer can recover and return to the admissible region.

All structural specifications share this likelihood implementation and are maximized with the Nelder-Mead simplex algorithm (Nelder and Mead, 1965), which requires no derivatives; where a smooth objective permits it, a quasi-Newton L-BFGS polish (Liu and Nocedal, 1989) follows. Before optimization, each coefficient is rescaled by the root-mean-square magnitude of its design-matrix column, which equalizes the local effect of a simplex step across an intercept, $\log n$, $[\log n]^2$, and standardized district characteristics; without this rescaling, the simplex spends most of its budget adapting to the very different natural scales of these regressors.

III.1.3 Profiling the leakage parameter

The parameter κ is estimated jointly with the model coefficients by maximum likelihood. Because the coefficient vector must be re-optimized at every candidate value of κ, we discretize the leakage parameter for computational tractability: the reported estimate maximizes the profile log likelihood

$$\ell_q(\kappa) = \max_{\beta_r, \beta_t, \beta_m} \sum_{i \in q} \log L_i^{(G)}(\beta_r, \beta_t, \beta_m; \kappa)$$

over a grid of 100 equally spaced values on $[0.01, 0.2]$, re-estimating the complete coefficient vector at each grid point from multiple starting values. A 95 percent set for κ inverts the likelihood-ratio statistic: it collects the grid points whose profile log likelihood lies within 1.92 points of the maximum, 1.92 being one half of the 95 percent critical value of a χ^2_1 distribution.

This procedure is applied to the full sample (Table 20) and, separately, to each electorate-size stratum, each time period, each alternative eligible-voter floor, the CCM replication sample, and both extended-covariate samples (Table 14 in Appendix II). The fixed-κ arrival-function checks in Table 15 are not re-profiled: under $f(n) = \sqrt{n}$ or $f(n) = 1$, the parameter κ has different units than under the baseline $f(n) = n$, so re-profiled values would not be comparable. Figure 14 reports the profile restricted to large elections, where the individual component matters most in absolute terms.

III.1.4 Bootstrap inference

Standard errors are computed by a block bootstrap whose resampling unit is the county, in the spirit of cluster-robust bootstrap inference. A replicate draws the original number of counties with replacement and includes every sampled election from each drawn county; a county drawn more than once contributes another copy of its full block. Clustering at the county level preserves the within-county dependence created by persistent organizational infrastructure—congregations, civic associations—and by repeated liquor elections in the same localities.

For the baseline tables, the procedure is run for the full sample and for each electorate-size stratum. For the time-period table, the grouped model is bootstrapped separately within each period; for the extended-covariate table, both the $m = 1$ and grouped specifications are bootstrapped on the full and first-referendum samples. A grouped-model bootstrap holds the relevant specification-specific $\hat{\kappa}_q$ fixed and warm-starts each replicate from that specification's profile optimum; we do not re-profile κ inside each draw, so the reported standard errors are conditional on the profiled leakage value, exactly as the baseline inference is. For each model and sample, we replicated 200 times. Table 10 in Appendix II reports the $m = 1$ benchmark estimates whose standard errors are produced by this procedure.

III.1.5 Simulation and common scoring of the alternative models

All simulation-based moments in the paper use fixed random seeds and draw the requested number of simulated elections at each observed district's fitted parameters. The $m = 1$ and grouped models are scored on their native count likelihoods, described in Section III.1.1. Each alternative model delivers fitted per-eligible-voter probabilities $(\hat{q}_{R,i}, \hat{q}_{L,i})$ and is placed on a common count scale with

the trinomial likelihood over (votes for, votes against, abstentions),

$$\begin{aligned}\ell_i^{(M)} = \log \frac{n_i!}{V_{R,i}!\, V_{L,i}!\, (n_i - V_{R,i} - V_{L,i})!} \\ + V_{R,i} \log \hat{q}_{R,i} + V_{L,i} \log \hat{q}_{L,i} \\ + (n_i - V_{R,i} - V_{L,i}) \log(1 - \hat{q}_{R,i} - \hat{q}_{L,i}),\end{aligned}$$

with eligible voters rounded to the nearest integer and probabilities floored at 10^{-12} so that no model receives an infinite penalty from a single observation. Because the native families differ—independent Poisson, compound Poisson, and trinomial—we complement the raw likelihoods with the Akaike and Bayesian information criteria, the county-clustered Vuong (1989) statistic, the distribution-free sign test of Clarke (2007), and the simulated turnout and margin moments reported in the main text.

Pivotal voters The pivotal-voter alternative follows the finite model of Coate et al. (2008). Nature assigns each of the n_i citizens to the supporting side with probability μ_i, citizens draw private voting costs from a uniform distribution, and each side turns out when its cost is below a cutoff determined by the probability of being pivotal. We retain the published index structure exactly. The preference share μ_i is a function of an intercept, the Baptist population share, and the MSA indicator. The two sides' benefit parameters are exponentials of linear indices in an intercept, the two proposition-type indicators (off-premise sale of all alcoholic beverages, and beer and wine only), the city-charter indicator, and the unincorporated-community indicator. The voting-cost scale is an exponential of the weekend-election indicator, for a total of fourteen parameters.

For each election and each candidate parameter vector, the two equilibrium cost cutoffs solve the pair of indifference conditions in which the marginal voter's cost equals her side's benefit times the probability that her vote is pivotal. We solve this system by damped fixed-point iteration, with up to 20,000 iterations at a tolerance of 10^{-7}. The pivot probability treats the vote difference among the other citizens as the difference of two Poisson counts, which follows the Skellam (1946) distribution; its Bessel-function expression is evaluated by a convergent series for small arguments and by an asymptotic expansion for large ones, so the computation remains stable in electorates of half a million voters. Estimation maximizes the trinomial count likelihood above by iterated Nelder–Mead restarts from the published coefficient vector: each restart runs the simplex to its iteration budget and restarts it from its own optimum, stopping when a complete restart improves the log likelihood by less than one point. At the reported estimates, the equilibrium conditions solve to tolerance in 96 percent of elections, with a maximum residual of 2.5×10^{-4}. The small-sample CCM replication in the main text is not re-estimated: there, we evaluate the published coefficients on the original sample instead of replicating the estimation.

Expressive voters In the expressive alternative, citizens vote for the expressive benefit of participation rather than because they may be pivotal. With $s_i = \sigma(x_i'\beta_s)$ the reform-support share and $\rho_{R,i} = \sigma(w_i'\beta_R)$, $\rho_{L,i} = \sigma(w_i'\beta_L)$ the side-specific participation propensities, the implied vote probabilities are $\hat{q}_{R,i} = s_i\rho_{R,i}$ and $\hat{q}_{L,i} = (1 - s_i)\rho_{L,i}$. Neither index contains the electorate size or any equilibrium object—that is the economic content of expressive voting. The support index x_i contains an intercept, the Baptist population share, and the MSA indicator; each participation index w_i contains an intercept, the two proposition-type indicators, the city-charter indicator, the unincorporated-community indicator, and the weekend indicator, for fifteen parameters in total. The parameters are estimated by maximum likelihood on the trinomial count likelihood, initialized at the logit of the sample mean vote share and turnout, using Nelder–Mead followed by a quasi-Newton polish; the reported fit converged.

Group rule-utilitarian voting The model of Coate and Conlin (2004) has a continuum of citizens. Conditional on the support share μ_i, turnout on each side is deterministic, so the model implies a closed-form joint density for the two vote shares, and the published estimator maximizes exactly that share density; there are no integer counts or sampling noise anywhere in the original procedure. We first replicate their estimator term for term: maximum likelihood on the share density with the published functional forms, in which the support share is Beta-distributed with shape parameters $\nu_i = \exp(\text{intercept, Baptist share, city})$ and $\omega = \exp(\text{intercept})$, the two sides' benefit parameters are exponentials of linear indices in an intercept, the proposition-type indicators, and the city indicator, and the voting-cost scale is an exponential of the weekend indicator—thirteen parameters. Estimation uses Nelder–Mead followed by a quasi-Newton polish, starting from the published coefficient vector; the vote shares of the eight elections with zero votes on one side receive a half-vote continuity correction so that the share density is well defined, and the reported fit converged.

To place the fitted model on the common count scale, we exploit the timing of the model's decisions. Because the critical cost levels of their Proposition 1 are chosen before the support share is realized, the equilibrium voting rates $(\hat{\sigma}_{R,i}, \hat{\sigma}_{L,i})$ of the two sides do not depend on μ_i, and the model's per-voter probabilities at any realization are $(\mu_i\hat{\sigma}_{R,i}, (1 - \mu_i)\hat{\sigma}_{L,i})$. The count likelihood therefore integrates the trinomial likelihood over the fitted Beta support-share distribution $\hat{F}_i$:

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

The integral is evaluated by equal-probability quadrature at the 101 quantiles of $\hat{F}_i$ corresponding to probabilities $(g - 0.5)/101$ for $g = 1, \ldots, 101$, with the terms aggregated in log space, and simulations redraw μ_i from the same fitted distribution in every replication. The common shock to

benefits and costs that enters their likelihood is held at its median, so no model in the comparison receives district-level unobserved heterogeneity beyond what its mechanism generates.

The Shachar–Nalebuff model The model of Shachar and Nalebuff (1999) identifies unobserved party-leader effort from the cross-state structure of presidential elections: followers' turnout responds to leader effort, and effort is pinned down by first-order conditions in which the value of winning a state combines its simulated Electoral College pivotality with delegate incentives, validated against survey data on campaign contact. A local liquor referendum has no analogue of these cross-state shifters and no observed measure of local campaign effort, so the model cannot be transported to our setting as published, and we exclude it from the comparison in Table 8.

For completeness, Table 21 reports estimates of a referendum adaptation that preserves the model's two central mechanisms while replacing the unidentified effort process with its equilibrium determinants. Followers behave as in their Assumption 1: a citizen votes when a uniform voting cost falls below an exponential participation index in the proposition-type, city, unincorporated-community, and weekend indicators. Leader effort is common to the two sides, as in their Proposition 1, and in their equilibrium it is a function of the predicted closeness of the race and of the population to be mobilized; we therefore proxy effort by these determinants, entering the predicted closeness of the referendum, $1 - |2\hat{d}_i - 1|$, and the centered logarithm of the electorate size in both sides' participation indices. The predicted support share $\hat{d}_i$ comes from a first-stage least-squares fit of the log-odds of the observed vote share on the Baptist share and the MSA indicator, mirroring the auxiliary closeness forecast that the original model feeds into its structural stage; the participation and effort parameters then maximize the trinomial count likelihood at the fitted support shares.

The adaptation reproduces the model's two signature comparative statics in our data: participation increases in predicted closeness (coefficient 0.53) and declines with electorate size (coefficient -0.22 on $\log n$). Its count log likelihood, $-1{,}187{,}419$, is in the same range as the other no-group alternatives and far below the grouped model. And because voting decisions remain conditionally independent across citizens, the adaptation inherits their central limitation: simulated margins of victory collapse relative to the data, with $\Pr(|M| > 0.10) = 0.030$ against an empirical frequency of 0.202.

Simulated moments and figure construction All simulated moments in the comparison table are computed from 1,000 simulated elections per observed election, drawn at each model's fitted election-specific probabilities with fixed random seeds; the grouped model simulates its block and residual components directly, and the integrated rule-utilitarian model redraws the support share

before each trinomial draw. The simulated density in Figure 10 of Appendix II.6 pools the simulated draws across elections and plots a Gaussian kernel density, with reflection at zero, against the histogram of the 1,703 observed outcomes. In these simulated distributions, only the integrated rule-utilitarian model generates visible margin dispersion—inherited from its fitted Beta support-share distribution—while the expressive and pivotal-voter models concentrate the margin in a spike at zero.

The election-level figures in the main text and in Appendix II.6 are constructed identically for every model: nine elections are selected at fixed quantiles of the electorate-size distribution (the 5th to the 99th percentile), 300 outcomes are simulated from each fitted model for each selected election, and the simulated (turnout, margin) pairs are plotted with the observed outcome marked by a star. Because the elections are pinned to fixed quantiles and the seeds are fixed, the panels are identical across models up to the model-specific draws, so the figures can be compared cell by cell.

III.1.6 Tables and Figures

Table 20: Compliance Parameter Estimate

	Full sample
$\hat{\kappa}$ (joint MLE)	0.0692
Log likelihood at $\hat{\kappa}$	−36813.8
95% likelihood-ratio interval	[0.0692, 0.0698]
Elections	1,703

Notes: The compliance parameter κ is estimated jointly with the grouped-model parameters by maximum likelihood. The 95% set inverts the profile likelihood-ratio statistic, $\{\kappa : \ell(\kappa) \geq \ell(\hat{\kappa}) - 1.92\}$, on the profiled grid; the profile re-estimates every grouped-model coefficient at each grid value of κ, so each point is the best of many starting values.

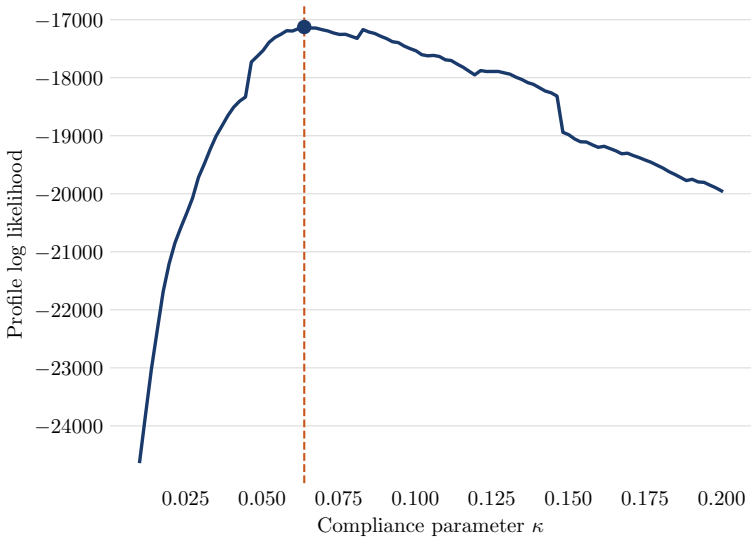


Figure 14: Large-election compliance diagnostics: the grouped-model likelihood profiled over κ for elections with $n \geq 8{,}500$, re-estimating all coefficients at each grid point.

Table 21: Follow-the-Leader Model Adapted to Referenda

Parameter	Estimate
Participation (for): intercept	−2.4221
Participation (for): alcohol (off-premise) proposition	−0.1480
Participation (for): alcohol (on/off-premise) proposition	−0.0388
Participation (for): city	0.9592
Participation (for): unincorporated	0.3225
Participation (for): weekend	−0.3402
Participation (against): intercept	−2.4063
Participation (against): alcohol (off-premise) proposition	−0.0704
Participation (against): alcohol (on/off-premise) proposition	0.0515
Participation (against): city	0.5999
Participation (against): unincorporated	0.6457
Participation (against): weekend	−0.2341
Effort: predicted closeness	0.4731
Effort: $\log(n)$ (centered)	−0.2209
Support share (first stage): intercept	0.0968
Support share (first stage): Baptist share	−0.5259
Support share (first stage): MSA	0.1002
N	1,703
Count log likelihood	−1196156.9
Simulated mean turnout	0.260
Simulated SD turnout	0.170
Simulated SD margin	0.045
Simulated $\Pr(\lvert M\rvert > 0.10)$	0.031

Notes: Estimates of the referendum adaptation of the follow-the-leader model. The support share is fitted in a first stage by least squares on the log-odds of the observed vote share, mirroring the auxiliary closeness forecast of the original model. Citizens vote when a uniform cost falls below an exponential participation index; equilibrium leader effort, common to the two sides, is proxied by its equilibrium determinants, the predicted closeness of the referendum, $1 - |2\hat{d} - 1|$, and the centered logarithm of the electorate size. The participation and effort parameters maximize the trinomial count likelihood at the fitted support shares. $M = (V_R - V_L)/n$; simulated moments use 1,000 simulated elections per observed election.

III.2 Omitted Proofs (Online appendix)

III.2.1 Omitted details in the proof of Proposition 3

We present here techical details omitted for brevity from the proof of Proposition 3 in the text. We first construct an exact interior equilibrium of the modified game in which independent voters abstain. Let $x_R = \frac{\alpha_R n}{m}\beta_R$, $x_L = \frac{\alpha_L n}{m}\beta_L$. These variables are the expected numbers of voting R- and L-groups. Let N_R and N_L be independent Poisson random variables with means x_R and x_L, respectively. The exact pivot probabilities of an R-group and an L-group are

$$\begin{aligned} P_R(x_R, x_L) &= \frac{1}{2}\Pr\{N_R = N_L\} + \frac{1}{2}\Pr\{N_R = N_L - 1\} \\ &= \frac{1}{2}\sum_{k=0}^{\infty} \frac{e^{-x_R}x_R^k}{k!}\frac{e^{-x_L}x_L^k}{k!}\left(1 + \frac{x_L}{k+1}\right), \qquad (45) \\ P_L(x_R, x_L) &= \frac{1}{2}\Pr\{N_R = N_L\} + \frac{1}{2}\Pr\{N_R = N_L + 1\} \\ &= \frac{1}{2}\sum_{k=0}^{\infty} \frac{e^{-x_R}x_R^k}{k!}\frac{e^{-x_L}x_L^k}{k!}\left(1 + \frac{x_R}{k+1}\right). \qquad (46) \end{aligned}$$

Both functions are continuous and strictly positive on $(0,\infty)^2$. Set $q = \frac{v_L}{v_R} \in (0,1)$ and $z = \sqrt{\frac{x_L}{x_R}}$. Writing $x_R = x$ and $x_L = z^2 x$, the exact formulas can be expressed using the modified Bessel functions of the first kind:

$$\begin{aligned} P_R(x, z^2x) &= \frac{1}{2}e^{-(1+z^2)x}\left[I_0(2zx) + zI_1(2zx)\right], \qquad (47) \\ P_L(x, z^2x) &= \frac{1}{2}e^{-(1+z^2)x}\left[I_0(2zx) + \frac{1}{z}I_1(2zx)\right]. \qquad (48) \end{aligned}$$

Define $A(x,z) = \frac{e^{-(1-z)^2x}}{4(\pi zx)^{1/2}}$ and, for $\nu = 0, 1$, $B_\nu(x,z) = (4\pi zx)^{1/2}e^{-2zx}I_\nu(2zx)$. Equations (47)–(48) become

$$\begin{aligned} P_R(x, z^2x) &= A(x,z)\left[B_0(x,z) + zB_1(x,z)\right], \qquad (49) \\ P_L(x, z^2x) &= A(x,z)\left[B_0(x,z) + \frac{1}{z}B_1(x,z)\right]. \qquad (50) \end{aligned}$$

For z in any compact subset of $(0,\infty)$, the standard large-argument expansion of the modified Bessel functions gives

$$B_0(x,z) = 1 + \frac{1}{16zx} + O(x^{-2}),\ B_1(x,z) = 1 - \frac{3}{16zx} + O(x^{-2}),$$

where the remainder terms and their first derivatives with respect to z are uniform on the compact set. In particular, $B_\nu(x,z) \longrightarrow 1$ and $\frac{\partial B_\nu(x,z)}{\partial z} \longrightarrow 0$ locally uniformly in z as $x \to \infty$. Define the normalized difference between the two groups' expected benefits:

$$G(x,z) = \frac{v_R P_R(x, z^2x) - v_L P_L(x, z^2x)}{A(x,z)}.$$

Since $A(x,z) > 0$, the equation $G(x,z) = 0$ is exactly the condition that the two group indifference equations be compatible. From (49)–(50),

$$G(x,z) = v_R \left[B_0(x,z) + zB_1(x,z)\right] - v_L \left[B_0(x,z) + \frac{1}{z}B_1(x,z)\right].$$

Consequently, $G(x,\cdot)$ and its derivative converge locally uniformly to $g(z) = v_R(1+z) - v_L\left(1+\frac{1}{z}\right)$. The function g is strictly increasing because $g'(z) = v_R + \frac{v_L}{z^2} > 0$, and its unique zero is $q = v_L/v_R$. Choose $\eta > 0$ such that $q - \eta > 0$. Then $g(q-\eta) < 0 < g(q+\eta)$ and $\inf_{z\in[q-\eta,q+\eta]} g'(z) > 0$. The locally uniform convergence of G and its derivative implies that there exists $\bar{x} < \infty$ such that, for every $x \geq \bar{x}$, $G(x, q-\eta) < 0 < G(x, q+\eta)$ and $\frac{\partial G(x,z)}{\partial z} > 0$ for all $z \in [q-\eta, q+\eta]$. It follows that, for every $x \geq \bar{x}$, there is a unique $\zeta(x) \in (q-\eta, q+\eta)$ such that $G(x,\zeta(x)) = 0$. Since the derivative with respect to z does not vanish at the root, the implicit-function theorem implies that ζ is locally continuous. Uniqueness makes the local branches coincide on overlaps, so they define a single continuous function $\zeta : [\bar{x},\infty) \longrightarrow (q-\eta, q+\eta)$. Moreover, $\zeta(x) \longrightarrow q$. Indeed, fix any $\varepsilon \in (0,\eta)$. Since $g(q-\varepsilon) < 0 < g(q+\varepsilon)$, local uniform convergence implies that, for all sufficiently large x, $G(x, q-\varepsilon) < 0 < G(x, q+\varepsilon)$. The strict monotonicity of $G(x,\cdot)$ then gives $q - \varepsilon < \zeta(x) < q + \varepsilon$. Thus, along the exact indifference locus, $\frac{x_L}{x_R} = \zeta(x_R)^2 \longrightarrow q^2 = \left(\frac{v_L}{v_R}\right)^2$.

Define the exact benefit of an R-group along this locus:

$$H(x) = v_R P_R\left(x, \zeta(x)^2 x\right), \qquad x \geq \bar{x}.$$

The function H is continuous and strictly positive. Furthermore, (49) and $\zeta(x) \to q < 1$ imply $H(x) \longrightarrow 0$ as $x \to \infty$. Fix $X \geq \bar{x}$. For every $c \in (0, H(X))$, there exists $Y > X$ such that $H(Y) < c$. Therefore, the intermediate-value theorem implies that the set $\mathcal{R}(c) = \{x \geq X : H(x) = c\}$ is nonempty. Define

$$x_R(c) = \min \mathcal{R}(c) \qquad \text{and} \qquad x_L(c) = \zeta(x_R(c))^2 x_R(c).$$

The minimum exists because $H(X) > c$, $H(Y) < c$, and the root set is closed. By construction, $c = v_R P_R(x_R(c), x_L(c)) = v_L P_L(x_R(c), x_L(c))$. Thus, $(x_R(c), x_L(c))$ satisfies both exact group indifference conditions.

We next show that $x_R(c) \longrightarrow \infty$ as $c \downarrow 0$. For every $M > X$, continuity and strict positivity imply $\min_{x\in[X,M]} H(x) > 0$. Hence, when c is smaller than this minimum, $\mathcal{R}(c)$ contains no point in $[X, M]$, and therefore $x_R(c) > M$. It follows that $x_R(c) \longrightarrow \infty$, $x_L(c) \longrightarrow \infty$, $\frac{x_L(c)}{x_R(c)} \longrightarrow \left(\frac{v_L}{v_R}\right)^2$.

For each sufficiently small c, define $N(c) = 1 + \left\lceil \max\left\{\frac{mx_R(c)}{\alpha_R}, \frac{mx_L(c)}{\alpha_L}\right\}\right\rceil$. For every $n \geq N(c)$, set $\beta_R^n(c) = \frac{mx_R(c)}{\alpha_R n}$, $\beta_L^n(c) = \frac{mx_L(c)}{\alpha_L n}$. These probabilities belong to $(0,1)$ and generate expected numbers of voting groups equal to $x_R(c)$ and $x_L(c)$. The exact indifference conditions above imply

that both types of groups are best responding. Hence, these strategies form an exact interior equilibrium of the modified game in which independent voters abstain.

We now verify that abstention is optimal for the independent voters. When all other independent voters abstain, an independent voter is pivotal only when the two numbers of voting groups are equal. Therefore,

$$P_I(x, z^2x) = \frac{1}{2}e^{-(1+z^2)x}I_0(2zx) = A(x,z)B_0(x,z).$$

Using the exact group indifference condition and $z = \zeta(x_R(c))$, we obtain

$$\begin{aligned} \frac{v_I P_I(x_R(c), x_L(c))}{c} &= \frac{v_I}{v_R}\frac{B_0(x_R(c), z)}{B_0(x_R(c), z) + zB_1(x_R(c), z)} \\ &\longrightarrow \frac{v_I}{v_R(1+q)} = \frac{v_I}{v_R + v_L} < 1. \end{aligned}$$

Thus, for all sufficiently small c, independent voters strictly prefer to abstain. The strategies constructed above are therefore equilibria of the original game.

Finally, consider the probability of the electoral outcome. Let $Z(c) = N_R - N_L$. Then $E[Z(c)] = x_R(c) - x_L(c)$, $\text{Var}(Z(c)) = x_R(c) + x_L(c)$. Because $x_L(c)/x_R(c) \to q^2 < 1$ and $x_R(c) \to \infty$, $\frac{x_R(c)-x_L(c)}{\sqrt{x_R(c)+x_L(c)}} \longrightarrow \infty$. The probability that L wins is no greater than $\Pr\{Z(c) \leq 0\}$. Chebyshev's inequality gives

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

Choose $c(\delta) > 0$ small enough that, for every $c \in (0, c(\delta))$, independent voters strictly prefer to abstain and the last probability is lower than δ. Then, for every such c and every $n \geq N(c)$, the strategies constructed above form an equilibrium in which independent voters abstain and L wins with probability lower than δ. ■

III.2.2 Omitted details in the proof of Proposition 4

If $v_I \leq v_L$, inequality (44) immediately rules out any equilibrium in which independent voters mix. Hence it suffices to consider $v_I \geq v_L$, and we write $\vartheta = v_I - v_L \in [0, \vartheta_0]$. Fix $\vartheta_0 > 0$ such that $v_I - v_L = \vartheta \in [0, \vartheta_0]$. In every equilibrium considered in Case 1, independent voters mix, and hence $\mathcal{P}_I(\sigma_n, \beta) = \frac{c}{v_I} \geq \frac{c}{v_L + \vartheta_0} > 0$. We proceed in two steps.

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 (43), $\mathcal{P}_I(\sigma_n, \beta) \leq \sup_{i\in\mathbb{Z}} S_i(T_R^n, T_L^n)$. 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 $\vartheta \in [0, \vartheta_0]$, there exists a finite constant $\overline{M}$ such that, for every n, every such ϑ, and every equilibrium in Case 1, $T_j^n \leq \overline{M} \qquad \text{for } j = I, L, R.$

Step 2. We next establish a uniform lower bound on T_R^n. We claim that there exist $\vartheta_1 > 0$, $N_1 < \infty$, and $\underline{M} > 0$ such that, for every $\vartheta \in [0, \vartheta_1]$, every $n \geq N_1$, and every equilibrium in Case 1, $T_R^n \geq \underline{M}$. Suppose otherwise. Then there would exist sequences $\vartheta_k \in [0, 1/k]$, $n_k \geq k$, and equilibria in Case 1 such that $T_R^{n_k} \longrightarrow 0$. Since Case 1 implies $T_R^{n_k} \geq T_L^{n_k}$, we also have $T_L^{n_k} \longrightarrow 0$. By the uniform upper bound, after passing to a subsequence, $T_I^{n_k} \to T_I^\infty$. Since $\vartheta_k \to 0$, equation (43) implies $\frac{c}{v_L} = \frac{1}{2}e^{-T_I^\infty}$, $T_I^\infty = \ln\left(\frac{v_L}{2c}\right)$. In the limit in which no other organized group votes, the pivot probability of an R-group is $\mathcal{P}_R^\infty = \frac{1}{2}p_0 + \sum_{j=1}^{m-1} p_j + \frac{1}{2}p_m$, where $p_j = \frac{(T_I^\infty)^j e^{-T_I^\infty}}{j!}$. Since $\frac{1}{2}p_0 = \frac{c}{v_L} > \frac{c}{v_R}$, we have $\mathcal{P}_R^\infty > \frac{c}{v_R}$. By continuity, the same strict inequality holds along the sequence for all sufficiently large k. But $T_R^{n_k} \to 0$ implies $\beta_R^{n_k} < 1$ eventually, so equilibrium requires $\mathcal{P}_R^{n_k} \leq \frac{c}{v_R}$, a contradiction. This proves the claimed uniform lower bound. ■