A Theory of Electoral Turnout Based on Free-Riding Aversion

Dan Bernhardt* Laurent Bouton† Stefan Krasa‡ Francesco Squintani§

July 5, 2025

Abstract

We propose a novel model of costly voting in a canonical spatial setting. Our main innovation is that citizens are averse to free riding on others. In particular, a citizen incurs a disutility when the benefits from others' voting efforts exceed her net cost of voting, but she, herself, abstains. We first establish the existence of equilibria that are robust to minimal coordination among voters, i.e., where a marginal small switch of a group of citizens from non-voting to voting cannot make them all strictly better off. We then derive how the underlying primitives of the electoral environment—e.g., voting costs, elite polarization, and the distributions of voter bliss points—affect equilibrium turnout and turnout rates. Finally, we endogenize platform choices and explore the consequences of these primitives for candidate platforms. Our comparative static predictions find support in the data, and we provide testable insights into how voter motives influence turnout, turnout rates, and candidate platforms.

*Department of Economics, University of Illinois, danber@illinois.edu, Department of Economics, University of Warwick, m.d.bernhardt@warwick.ac.uk

†Department of Economics, Georgetown University and NBER. lb910@georgetown.edu

‡Department of Economics, University of Illinois, skrasa@illinois.edu

§Department of Economics, University of Warwick, f.squintani@warwick.ac.uk

1 Introduction

Voters are at the heart of the democratic decision-making process. They influence policy directly through referenda and ballot initiatives, and indirectly by electing representatives and holding them accountable. In 2024 alone, approximately 70 nations—comprising roughly half of the world's population—held elections to choose their national leaders, including presidents and national assembly members. Beyond these national-level elections, citizens often participate in local elections to select state legislators, mayors, school board members, etc. Despite this, scholars remain divided on how to model voter behavior, and existing models deliver predictions at odds with empirical evidence. This impairs our understanding of voting behavior and our ability to address policy questions. Moreover, even when a political economy model does not explicitly focus on voters, accurately accounting for voter behavior is essential to capture voters' influence on other actors and dimensions of the democratic decision-making process.

The lack of consensus does not imply an absence of voting models. To the contrary, many model variants exist, including the so-called expressive voting, pivotal voting, and group-based voting models. However, each of these models has serious weaknesses. As we elaborate in the literature review, the expressive voting and pivotal voting models both fail to align with empirical evidence: expressive voters are unrealistically insensitive to their electoral environment, while pivotal voting probabilities almost vanish in large elections, providing essentially no reason to vote, leading to the *paradox of voting*. Group-based models avoid these pitfalls, as they can predict high voter turnout that is responsive to the characteristics of the electoral environment. However, to deliver such predictions, these models exogenously divide the electorate into large groups or coalitions. The lack of criteria to determine which groups form makes it impossible to determine in which situations, if any, such large groups should form, much less how the sizes of such groups and electoral outcomes should vary with the underlying primitives.

This paper develops a new, tractable model of voting within the canonical spatial setting pioneered by Downs (1957). The key novel ingredient is that citizens are averse to free riding, which is modeled as a disutility of abstaining that increases in the benefit derived from the costly voting efforts of others. Our assumption of free-riding aversion is grounded in the behavioral and experimental literature. It is consistent with the observation first made in Kelley and Stahelski (1970) and subsequently confirmed by a large literature in experimental economics (e.g., Fischbacher and Gächter (2010); Fehr and Schurtenberger (2018)) that people are more cooperative if they believe that others will also be cooperative, and that they are more selfish if they believe that others will also be selfish. We are agnostic as to the reasons underlying why people are "conditional cooperators"—which the literature argues can be either "internal"reasons (e.g., altruism, civic duty, or ethical concerns) or "external" (related to social considerations such as social pressure, social image, or social norms). Recent evidence about the relevance of those internal and external factors for turnout decisions adds credence to our approach. We discuss this further in Section 1.2.

We consider a canonical spatial setting with two candidates, where voting is costly. There is a continuum of two types of citizens, non-partisans and partisans. The number of partisans is random from the viewpoint of non-partisans. Partisans always vote for their preferred candidate, regardless of platforms and costs. The utility of a non-partisan has two (weighted) components: A private expressive component from voting for their preferred candidate, and a term that captures the positive externality generated by other non-partisans who are voting for a candidate even though their costs of voting exceed their private expressive benefit. Non-partisans who choose not to vote incur a free-riding disutility that is proportional to this externality. When deciding whether or not to vote, non-partisans take the voting behavior of others as given, and compare their private expressive benefits and costs of voting with the disutility of free-riding. In the Nash equilibrium, the "assumed" voting behavior of other citizens corresponds to their actual, optimal, choices.

It is immediate that a Nash equilibrium always exists in which either nobody votes, or only citizens driven by expressive motives vote. In such an equilibrium, no citizen feels to be free riding as no other citizen makes a *costly* voting effort. We then characterize Nash equilibria of our model with greater voter participation that satisfy a no minimal expansion condition. That is, we require that it is not possible for an arbitrarily small group of non-voters to switch to voting thereby making themselves all strictly better off.

We prove existence of a "no minimal expansion" equilibrium in symmetric settings. To facilitate further characterizations we assume that the distribution of the random net partisan vote is uniform. When this is so, the behavior of non-partisans on the left is unaffected by the behavior of non-partisans on the right, and vice versa. We characterize the unique "no marginal expansion" equilibrium for two boundary scenarios: when non-partisans are purely selfish (i.e., they have expressive voting considerations, and no free-riding ones), and when they are purely cooperative (i.e. they have free-riding considerations, and no expressive voting ones). These boundary scenarios help us highlight the qualitative consequences of the nature of citizen motives, and to understand the mixed motive setting, whose equilibrium properties prove to be a weighted combination of these two polar opposite scenarios.

With selfishly-motivated non-partisans, incentives to abstain are the same for moderates and non-moderates. Selfish non-partisans on both sides of a candidate's platform can be "alienated." For cooperatively-motivated non-partisans, the situation is different: while moderates may abstain because they care little about who wins, non-moderates are never alienated. This latter result reflects that non-moderates (i) are more reactive to differences in platforms that are further away from their bliss point, and (ii) do not bear an expressive cost of voting for a candidate whose policy platform is far from their preferred policy. The result that non-moderates are never alienated is the extreme manifestation of a feature that holds quite generally in our model—with cooperative motives, moderates are less willing to participate than non-moderates.

We then derive how varying the primitives of the electoral environment impacts non-partisan turnout. We focus on the comparative statics of turnout that speak to empirical studies identifying a clear causal

impact of specific features of the electoral environment.

We show that, consistent with empirical findings and regardless of the nature of non-partisan preferences, higher voting costs reduce non-partisan turnout. In contrast, several other primitives have sharply different effects on turnout, according to whether cooperative or selfish preferences dominate. For example, increasing elite polarization, in the form of a symmetric spreading of candidate platforms, increases turnout if non-partisans have purely cooperative preferences, but does the opposite if non-partisans have purely selfish preferences. With cooperative preferences, the intuition for the increased turnout reflects that shifting platforms further apart raises the stakes of winning the election. With selfish preferences, the reduced turnout induced by increased polarization reflects that the set of non-partisans close enough to the candidate's platform for whom the expressive benefits of voting outweigh the cost is shifted toward the tail where there are fewer non-partisans.

We also explore the impact of differences in the electoral support for the two candidates on turnout rates. Theoretical research based on pivotal voting considerations predicts an *underdog effect*—supporters of the underdog should turn out at a higher rate than supporters of the front-runner. This prediction has been tested in the laboratory and most studies (Duffy and Tavits (2008); Großer and Schram (2010); Kartal (2015); Agranov et al. (2018)) find the opposite result: supporters of the front-runner turn out at a *higher rate*—the so-called *bandwagon effect*. Our model predicts that a bandwagon effect should emerge if cooperative motives are sufficiently strong. Indeed, cooperatively-motivated non-partisans are subject to a discouragement effect when they anticipate having fewer co-supporters, as the benefit from the costly voting effort of others then decreases; and to an encouragement effect when they anticipate having more co-supporters.

We conclude our analysis by endogenizing candidate platform choices along the lines of Wittman (1983) and Calvert (1985) within our voting model. We first show that as long as voting costs are not too high, elite polarization (i.e., the distance between the platforms of the two candidates) is greater when non-partisans have purely selfish preferences than when they have purely cooperative preferences. To understand the intuition, remember that with selfish preferences only non-partisans sufficiently close to a candidate's platform participate. This means that while moving a platform outward toward the candidate's ideal policy alienates some moderates, it also convinces some non-moderates to vote. Hence, the cost to a candidate of moving closer to his ideal policy is limited. By contrast, with cooperative preferences, only moderates abstain. Hence, a move to the extreme may push more moderates to abstain, but without the benefit of convincing more non-moderates to vote. The impact on the probability of winning is thus larger, reducing the incentives of candidates to move outward toward their ideal policies.

Finally, we explore the impact of voting costs on elite polarization. With purely selfish preferences, higher voting costs increase polarization. This is because, when voting costs are higher, the marginal moderate and non-moderate voters are closer to the candidate's platform. With a single-peaked density over voter

ideal points, this reduces the net number of voters lost by choosing a more extreme platform, reducing the cost of locating more extremely. In contrast, when non-partisans' cooperative preferences are sufficiently strong, elite polarization *decreases* in voting costs. This reflects that higher voting costs reduce voting by moderates, i.e., the cutoff voters on each side become more extreme. As voters are risk averse, this makes the left candidate's marginal coalition member more sensitive to changes in the right candidate's platform than the right candidate's marginal coalition member. In particular, if the right candidate locates more moderately, he always gains more votes on the left than he may lose on the right. Hence, if the right candidate moves marginally to the center he gains more votes after the cost increase, incentivizing moderation.

1.1 Literature

There is a large literature seeking to understand the drivers of voter turnout. The classic model of strategic voting considers fully rational and instrumental agents, who turn out to vote only if the probability that their vote is pivotal in determining the winner of the election is sufficiently large relative to the cost of voting (Riker and Ordeshook (1968), Palfrey and Rosenthal (1983), Palfrey and Rosenthal (1985), Myerson (1998), Castanheira (2003), Myatt 2012). The pivotal voter model is afflicted by a serious weakness: the vanishingly low turnout levels it predicts in large elections—highlighting the *paradox of voting*.[1,2] A common alternative approach, the expressive voting model, assumes that voters enjoy the act of voting for their preferred candidate, which they compare to their cost of voting when deciding whether to participate. While research has shown that voters derive a direct "expressive" benefit from voting for a candidate whose platform closely aligns with their preferences,[3] this model's predictions are also at odds with the facts.[4]

[1]Models with aggregate uncertainty about the distribution of preferences predict a turnout rate that converges to zero at a much slower rate than models without aggregate uncertainty (Good and Mayer (1975), Castanheira (2003), Myatt 2012). Nonetheless, for large electorates, these models still predict counterfactually low turnout rate.

[2]Another weakness of the pivotal voter model is its predictions of the behavior of voters in elections with more than two candidates. See the discussions in, e.g., Fey (1997), Bouton et al. (2015), or Bouton and Ogden (2021).

[3]There is clear evidence that voters receive a direct payoff from supporting a particular candidate while also caring about the election outcome. Empirical studies based on election data (e.g., Fujiwara (2011); Pons and Tricaud (2018); Spenkuch (2018)) and laboratory experiments (e.g., Forsythe et al. (1996); Duffy and Tavits (2008); Esponda and Vespa (2014); Bouton et al. (2015)) support this dual motivation. The original concept of expressive voting was introduced by Brennan and Lomasky (1993).

[4]Several pieces of evidence are at odds with the predictions of the expressive voting model. First, in elections with more than two candidates, the literature comparing each voter's actual vote to her preferences finds that many voters cast misaligned votes—voting for a candidate other than their most-preferred one—see, e.g., Alvarez and Nagler 2000, Blais, Nadeau, Gidengil, and Nevitte 2001 and papers cited therein. Abramson et al. (2018) finds that the incidence of voting for a second-choice candidate was almost 40% in some constituencies for the 2010 British general election, while Daoust (2018) finds that 22.6% of voters selected their second-choice in the 2015 Canadian general election. Some studies (Kawai and Watanabe 2013, Spenkuch (2018)) find that a larger fraction of voters are strategic (estimates range from 40% to 85%), but they do not necessarily cast a misaligned vote. Finally, the experimental literature finds evidence of strategic voting both in two-candidate elections (for the costly voting side, see e.g., Levine and Palfrey (2005); for the Condorcet jury theorem side, see e.g., Guarnaschelli et al. 2000, Battaglini et al. 2008, 2010,

In reaction to the weaknesses of those widely-used models, researchers have proposed group-based voting models. There are three main variants of those models: (i) ethical voter models, which assume that individuals are ethical, or rule-utilitarian, in the sense that they decide whether to vote depending on what is optimal for their group as a whole or for the entire electorate (Harsanyi (1977), Feddersen and Sandroni (2006), Coate and Conlin (2004));[5] (ii) mobilization voting models, which assume that individuals' turnout decisions depend on how much mobilization effort leaders exert (Uhlaner (1989), Morton (1991), Shachar and Nalebuff (1999)); and (iii) social pressure models, which assume that individuals' decisions to vote are influenced by the pressure of their peers, e.g., they want to appear prosocial to others or avoid being punished by their peers (Ali and Lin (2013), Levine and Mattozzi (2020)). All these variants predict high turnout levels (and hence provide an explanation for the paradox of voting), and produce comparative statics that are empirically sound. While direct tests of those models on observational data are encouraging relative to both the purely expressive and pivotal voter models (Coate and Conlin, 2004; Coate et al., 2008), their absolute performance, even in the small elections being analyzed, is not fully satisfying.[6]

A more fundamental issue with group-based voting models is their assumption that group membership is exogenous. The rule utilitarian approach, by its nature, relies on groups existing ex-ante, as ethical rules have to be formulated before individuals learn their types.[7] For instance, in the ethical voting models of Feddersen and Sandroni (2006) or Coate and Conlin (2004), voters are grouped by the candidate whom they prefer and only consider the welfare of those voters.[8] Similarly, in Shachar and Nalebuff (1999), leaders can target mobilization effort only at their supporters. However, those models neither provide foundations for how those groups are formed, nor how the sizes of such groups and electoral outcomes vary with the underlying primitives. This is especially problematic because the ability of those models to escape the paradox of voting requires that the electorate consists of a sufficiently small number of such ex-ante determined groups.

In our approach, citizens do not belong to an exogenously defined group: they maximize their own utility

Goeree and Yariv 2011, Bhattachary et al. 2014, Bouton et al. 2017a) and in multi-candidate elections (see e.g. Forsythe et al. 1993 and 1996, Fisher and Myatt 2001, Morton and Rietz 2008, Bouton et al. 2016, 2017b).

[5]Roemer (2010, 2015) propose an alternative approach, which is not focused on voting. He assumes that agents have traditional preferences, but that the way they optimize their behavior deviates from the Nash equilibrium approach. In particular, he assumes that agents look for their best response considering that others would behave similarly, terming this a "Kantian equilibrium".

[6]Not only do group-based voting models poorly describe the data, as indicated by their low likelihood scores, but they fail to generate key patterns in the data, such as large victory margins in races with high turnouts. Merlo and Palfrey (2018) also document sub-par performances of ethical voter models in terms of concealed parameter recovery (a form of external validation analysis).

[7]An exception is Herrera and Martinelli (2006), who consider a model in which group leaders emerge endogenously out of the population and then compete for followers. Note however that there remains an exogenous dimension to the groups in this model, as leaders are exogenously and randomly assigned a group of voters that they can influence. Their effort then determines the share of individuals in their group that turns out to vote.

[8]As Harsanyi (1977) observes, "morally right acts ... maximize (ex-ante) social utility." Concretely, questions such as "is increasing taxes on the wealthy morally right?", can only be addressed ex-ante, prior to individuals learning their actual incomes.

function, taking the voting decision of all other citizens as given, as in any Nash equilibrium. Nonetheless, their voting behavior is influenced by the voting efforts of other citizens through their aversion to free riding. As a result, the set of citizen influencing the behavior of any given citizen—the "group" of a citizen—is determined endogenously, depending on whether and on whom a citizen feels to be free riding.[9] Our model thus predicts which citizens will join together to vote for a candidate, and under which conditions these groups are large enough to guarantee a high turnout rate.

This approach permits us to derive how voting groups and electoral outcomes vary with primitives. In particular, our model delivers comparative statics that qualitatively align with systematic patterns in voting behavior identified in the empirical literature. The ones we deem most important, due to their prominence in the literature, concern the impact on turnout of a change in (i) the cost of voting, (ii) elite polarization or the stake of the election, (iii) the platform of one of the two candidates, and (iv) how cooperative citizens are. We discuss the relevant empirical literature further, as we present the results. We also explore the impact of features of the electoral environment on the turnout rates of the supporters of the two candidates. As mentioned above, our finding that a bandwagon effect exists when voters are sufficiently cooperative, which is in stark contrast with the typical prediction of an underdog effect by extant models in the literature, can reconcile otherwise puzzling empirical findings.

Finally, our analysis is related to the literature on the strategic choice of platforms by candidates (Downs 1957, Wittman (1983), Calvert (1985), Lindbeck and Weibull 1987, Dixit and Londregan 1996, Persson and Tabellini 1999, 2000, Grossman and Helpman 2005, and Stromberg 2008, Bierbrauer et al.2022). As we assume a continuum of citizens with spatially differentiated bliss points, our model is more closely related to the "spatial politics" strand of that literature (Downs, 1957; Calvert, 1985; Wittman, 1983; Ledyard, 1984; Adams and Merrill, 2003). Most such studies assume that voters are expressive (see, e.g., Calvert (1985); Wittman (1983) for models with costless voting, and Adams and Merrill (2003), for costly voting).[10] These models predict full convergence to the bliss point of the (expected) median when candidates are purely office-motivated, and partial convergence when candidates are policy-motivated. Our model predicts that candidates adopt platforms closer to their bliss points when voters are selfish than when they are cooperative. Our model also produces substantially different comparative statics. For example, if voters' cooperative motives are sufficiently strong, increased voting costs may actually *reduce* elite polarization (the distance between equilibrium platforms). This has important implications for the evaluation of policies influencing the cost of voting. For instance, Oprea et al. (2024) argue, based on traditional models of spatial politics with expressive voters, that an important argument in favor of compulsory voting (a policy that increases the cost of not voting) is that it should decrease elite polarization. In our model, this conclusion can be reversed.

[9] See Bernhardt et al. (2024) for a model that applies these concepts to multi-candidate setting where voting is costless.

[10] An exception is Ledyard (1984), which embeds the pivotal voter model in a model of spatial politics.

1.2 Discussion of Free-Riding Aversion

A large body of evidence indicates that people free ride far less than what standard economic theory predicts. The experimental literature finds that a substantial share of individuals do not free-ride even when, from a rational and selfish perspective, it is a dominant strategy to do so (see the reviews by Gachter and Herrmann (2009); Fehr and Schurtenberger (2018)).[11] There are also many examples of real-life situations in which someone's behavior appears similarly "prosocial", i.e., irrational from a purely selfish perspective. For instance, millions routinely donate blood even if they receive no monetary reward, and incur discomfort; people often recycle, reduce water usage, or pay premiums for sustainable products, even when the personal benefits are tiny and the actions only meaningfully help if many others also cooperate; people often donate or volunteer to help victims of natural disasters with no expectation of personal gain; many people report income and pay taxes even when the risk of audit is low (see also Roemer (2015)). As Gachter and Herrmann (2009, p. 791) write, "...humans often manage to avoid the tragedy of the commons and achieve high levels of cooperation. This holds for hunter-gatherer societies to complex modern nation states ..."

Various explanations and theories have been proposed to explain these striking departures from standard economic theory. As Malmendier et al. (2014) explain, these theories can be classified according to the "internal" or "external" nature of the factors underlying the prosocial behavior. Internal factors relate to individual preferences, and encompass altruism, warm-glow, and inequity aversion, as well as the reciprocity they may generate (see, e.g., Andreoni and Miller (2002); Charness and Rabin (2002); Falk and Fischbacher (2006); Fehr and Schmidt (1999); Bolton and Ockenfels (2000); Levine (1998); Rotemberg (2008); Battigalli and Dufwenberg (2009); Rabin (1993)). External factors relate to social considerations such as social pressure and punishment, social image and norms, including reciprocity norms (see, e.g., Bénabou and Tirole (2006); Battigalli and Dufwenberg (2007); Andreoni and Bernheim (2009); Ali and Lin (2013); Fehr and Gächter (2002)).[12] These factors can cause people to behave as if they were averse to free riding.

Our approach captures the concept of "conditional cooperation" introduced by Kelley and Stahelski (1970): people are prosocial/cooperative if they believe that others will also be prosocial/cooperative, and they are selfish if they believe that others will also be selfish. Using the "strategy method," Fischbacher et al. (2001) provide evidence that many subjects are conditional cooperators, as they cooperate at high levels if they believe others are doing the same, but free ride if they believe others are not cooperating (see also Kocher et al. (2008) and Chaudhuri (2011)). Fischbacher and Gächter (2010) and Fehr and Schurtenberger (2018) observe that this concept can explain not only "excessive" cooperation, but also other patterns of

[11] These results are established in public good experiments where players simultaneously decide how much of their private endowment to contribute to a public good and how much to keep for themselves, creating a social dilemma between the individually-rational incentive to free-ride (contribute nothing) and the collectively-optimal outcome of full contribution by everyone.

[12] For experimental evidence of the importance of social considerations for prosocial behavior, see Fehr and Gächter (2000); Fehr and Gächter (2002); Cubitt et al. (2011); Kimbrough and Vostroknutov (2016), and references therein.

behavior uncovered by the experimental literature that other theories cannot.[13]

Our model embraces the concept of conditional cooperation and the fact that conditional cooperation is inherently an equilibrium concept: people's desire to cooperate depends on what they believe others will do. We capture this in a simple way via free-riding aversion: citizens get a disutility of abstaining that is increasing in how many other citizens vote for their preferred candidate. We remain agnostic as to the factors underlying free riding aversion and hence conditional cooperation. In line with the literature, conditional cooperation can be a motivation on its own, a consequence of fairness preferences or social considerations, or a combination of these factors. Existing attempts at formalizing conditional cooperation (e.g., Ambrus and Pathak (2011)) build on the literature on reciprocity (see Malmendier et al. (2014)). To the best of our knowledge, we are the first to formalize conditional cooperation through free-riding aversion and, more importantly, to apply the concept to turnout decisions and, more generally, to "group contest" settings.[14]

The literature provides evidence that both internal and external factors motivate citizens to behave pro-socially when deciding whether to vote. For internal factors, many studies explore the motives of voters using survey data (see, e.g., Blais (2000); Markus (1988); Knack (1992); Opp (2001); Jankowski (2007)). That literature highlights altruism, civic duty and ethical concerns as key motives for turnout. Fowler (2006) uses a laboratory experiment to show that altruism plays a significant role in turnout decisions, while Feddersen et al. (2009) use a laboratory experiment to provide evidence that expressive ethical concerns are highly relevant in large elections.

For external factors, Bond et al. (2012) and Gerber et al. (2008, 2016, 2017) use mail experiments to provide clear evidence of the relevance of social considerations. They find that participation increases substantially following information about the behavior of one's neighbors or promises to share information about one's turnout to their neighbors. DellaVigna et al. (2016) use a field experiment to provide direct evidence that citizens vote to tell others (or not to be forced either to admit that they abstained, or that they lied about it), and more generally that image concerns matter for voting behavior.[15] Funk (2010) uses a natural experiment in Switzerland following the introduction of optional postal voting to provide evidence about the relevance of social motives. She finds that, in small communities where social incentives are strong, turnout *fell* following the introduction of optional postal voting even though voting costs decreased, but it did not change in large communities (where social incentives are arguably weaker).

Finally, the literature provides direct evidence of the relevance of conditional cooperation in voting. Gerber and Rogers (2009) use get-out-the-vote field experiments to show that messages emphasizing high

[13] In particular, the facts that (i) free-riding increases over time when the public goods game is played repeatedly by the same set of players, and (ii) "excessive" cooperation resumes after a restart of the game with the same set of players.

[14] Kiss et al. (2020) provide experimental evidence that conditional cooperation is also important in group contests, and helps explain the consistent "over-expenditure" pattern uncovered in the experimental literature.

[15] Silver et al. (1986) and Belli et al. (1999) also find that a substantial share of abstainers lie about having abstained when asked.

expected turnout by other citizens are more effective at increasing turnout than messages emphasizing low expected turnout. These findings align with the core idea of conditional cooperation that people's desire to cooperate depends on what they believe others will do. Doherty et al. (2017) uses survey experiments to explore how the expected behavior of others affects one's decision to turnout. They find clear evidence of a positive relationship between the social benefit of voting and the expected voting behavior of others.

2 Model

There are two candidates, $i = L, R$, with policy positions x_i, and a continuum of citizens. There are two types of citizens—partisans and non-partisans. A partisan for candidate i always votes for i regardless of the policy positions taken. Thus, the partisan support for a candidate i is summarized by the number $\rho_i \geq 0$ of i's partisan supporters. Without loss of generality, we restrict attention to the distribution, G, over the net partisan turnout for candidate R, $\rho = \rho_R - \rho_L$. We assume that G has a symmetric density around 0.

Non-partisans are distinguished by their ideal policy bliss points, y, which we refer to as the non-partisan's type. Let Φ be the cdf over voter types y in $\mathbb{R}$, which is single-peaked and log-concave. With slight abuse of notation we also use Φ to denote the probability measure.

Let S_i, $i = L, R$, be the set of non-partisans who vote for candidate i. The utility of a non-partisan y who votes is comprised of two parts. First, non-partisans incur a cost $c > 0$ from voting. The second part measures the expressive utility from voting for a particular candidate, taking the form $A - (y - x)^2$, where x is the policy of the candidate for whom y votes, and $A \geq 0$ is a fixed positive expressive benefit of voting. Thus, the utility from voting for candidate i is[16]

$$u_y(i) = \beta(A - (y - x)^2) - c,$$

where $\beta \in [0, 1]$ is the relative weight citizen y places on his expressive utility.[17] The main innovative feature of our model is in the utility of a non-partisan who abstains: we assume that they dislike free-riding on the costly voting effort of others. In particular, citizen y experiences a disutility from abstaining proportional to the instrumental benefit y derives from the voting efforts of non-partisans who also vote for y's preferred candidate. We denote that free-riding disutility by $B_y(i, S_L, S_R)$, which we formally define below. The utility non-partisan y realizes from abstaining is thus:

$$u_y(N) = -(1 - \beta) B_y(i, S_L, S_R),$$

[16]The quadratic form facilitates explicit solution. Qualitatively, our analysis only hinges on non-partisans being globally strictly risk averse.

[17]We could add a third, instrumental part to the utility of non-partisans, who would then care about the identity of who wins the election. However, the impact of any individual vote on the electoral outcome is null due to the continuum of non-partisans, so this part would disappear when considering the utility from voting.

where $1-\beta$ is the weight on the free-riding disutility in preferences. We refer to $\beta = 1$ as the purely selfish preferences case, and to $\beta = 0$ as the purely cooperative preferences case.

We now derive the free-riding disutility function, $B_y(i, S_L, S_R)$. This function compares the utility of (non-partisan) citizen y if coalition S_i forms to his utility if the coalition does not form. Given that non-partisans with a sufficiently large expressive utility always vote, the minimal coalition in support of candidate i is comprised of those non-partisans. Let

$$E_i = \Big\{ y | -(y - x_i)^2 \geq -(y - x_{-i})^2, \text{ and } \beta(A - (y - x)^2) - c \geq 0 \Big\},$$

be the set of non-partisans who would vote for candidate i even if the free-riding disutility $B_y(i, S_L, S_R)$ were zero. That is, citizens in E_i would vote for candidate i for expressive reasons regardless of the voting behavior of other citizens. The instrumental benefit that citizen y derives from the voting efforts of other citizens is then given by y's utility gain if everyone in S_i actually votes for candidate i compared to the case where candidate i only receives the support of citizens in E_i. Thus, y's free-riding disutility equals

$$B_y(i, S_L, S_R) = \Big(P(S_L, S_R) - P(E_i, S_{-i}) \Big) \Big((x_R - y)^2 - (x_L - y)^2 \Big),$$

where $P(S_L, S_R) = G(\Phi(S_L) - \Phi(S_R))$ is the probability that candidate L wins given S_i, $i = L, R$, and $P(E_i, S_{-i})$ is the probability that candidate L wins if coalition S_i is replaced by E_i.

An equilibrium in our game is a pair of coalitions (S_L, S_R) such that every non-partisan y is best responding to the behavior of other non-partisans:

Definition 1 *(S_L, S_R) is a voting equilibrium if and only if all non-partisans maximize their utility given S_L and S_R, i.e., $y \in S_i$ if and only if $u_y(i) \geq u_y(N)$, and $y \notin S_L \cup S_R$ if and only if $u_y(N) > u_y(i)$, for $i = L, R$.*

It follows immediately that (E_L, E_R) is always a voting equilibrium (even if these sets are empty). In particular, the free-riding disutility $B_y(i, E_L, E_R) = 0$, and hence citizen optimization implies that only those in E_i vote for candidate i. However, starting from such an equilibrium, and marginally expanding the coalition of non-partisans may make every new member better off. In such a case it should be possible for the coalition to grow by reaching out to convince these individuals to join. This leads us to focus on equilibria that are robust to a minimal amount of coordination. In particular, we require equilibria to satisfy the following additional condition:

Condition 1 (No Minimal Expansion) *A voting equilibrium (S_L, S_R) violates no minimal expansion if and only if for all $\varepsilon > 0$ there exists a set of voters $T \subset \mathbb{R}$ with $\Phi(T) < \varepsilon$, and a candidate i such that all members of T are strictly better off voting for candidate i rather than abstaining.*

The no minimal expansion condition requires that a marginal expansion of the voting coalitions is not possible because some of the new members would be strictly better off abstaining instead of voting, i.e., they would strictly prefer not joining the expanded coalition. Henceforth, we refer to equilibria that satisfy the no minimal expansion requirement as "equilibria".

3 Equilibrium Analysis

We organize the equilibrium analysis in five parts. First, we prove preliminary results about some necessary features of equilibrium voting coalitions. These preliminary results are key to proving and understanding our main results. Second, we identify different sets of conditions that guarantee the existence of an equilibrium. Third, we characterize the unique equilibrium for two extreme, but relevant cases: purely selfish preferences ($\beta = 1$) and purely cooperative preferences ($\beta = 0$). The realistic setting with mixed motives, i.e., $\beta \in (0, 1)$, largely inherits the properties from these two extreme scenarios. Fourth, we prove some comparative static results about how different features of the electoral environment affect turnout. Finally, we explore the impact of these features on turnout rates, with a special focus on so-called bandwagon and underdog effects.

3.1 Preliminary Results

Citizen y's net benefit from voting for candidate i versus not voting is given by

$$\begin{aligned}\gamma_i(y, S_L, S_R) &= u_y(i) - u_y(N) \\ &= (1-\beta)\left(P(S_L, S_R) - P(E_i, S_{-i})\right)\left((x_R - y)^2 - (x_L - y)^2\right) + \beta\left(A - (y - x_i)^2\right) - c. \quad (1)\end{aligned}$$

Because B_y is affine linear in a citizen's ideal policy, y, and the expressive utility part is quadratic, it follows that γ_i is strictly concave in y. For $|y|$ sufficiently large, the quadratic term dominates the affine linear free-riding disutility, and hence the ideologically most extreme citizens do not participate when $\beta > 0$.

Next, note that, for a given citizen y, expanding the coalition to which she belongs increases her net benefit $\gamma_i(y, S_L, S_R)$. This is because the free-riding disutility is larger when the coalition is larger.

These observations underlie the following lemma:

Lemma 1 *Let (S_L, S_R) be an equilibrium. Then:*

1. *For $\beta > 0$, coalition S_i, $i = L, R$, is a closed interval with $S_i = [\underline{y}_i, \bar{y}_i] \supset E_i$.*
2. *$\bar{y}_L < (x_L + x_R)/2 < \underline{y}_R$.*
3. *If y is on the boundary of S_i then $\gamma_i(y, S_L, S_R) = 0$.*

The results in Lemma 1 stem from three facts: (i) $\gamma_i(y, S_L, S_R)$ is continuous and strictly concave in y, (ii) any citizen y with $\gamma_i(y, S_L, S_R) \geq 0$ must be part of coalition S_i, and (iii) any citizen y with $\gamma_i(y, S_L, S_R) < 0$, $i \in \{L, R\}$ cannot be part of a voting coalition. Figure 1 illustrates the content of Lemma 1 by depicting $u_y(L)$, $u_y(R)$, and $u_y(N)$, as well as the coalitions S_L and S_R.

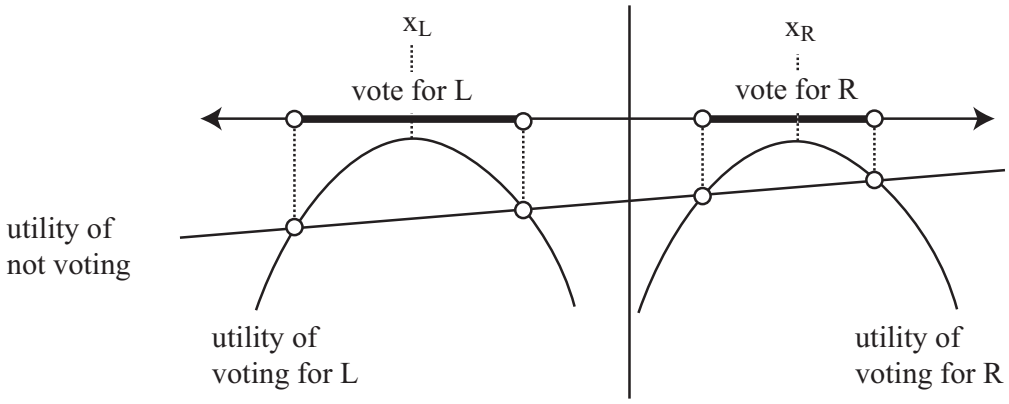


Figure 1: The horizontal axis is the ideological space. The vertical axis, positioned at the mid-point between x_L and x_R, gives the utility of different actions (voting for L, for R, or not voting). The subset of the ideological space in bold represent the citizens included in the coalitions S_L and S_R (with all citizens with positive utility of voting for L or R included in those coalitions, and all other not included).

We next consider the impact of expanding the coalition on the net benefit of its marginal member.

Lemma 2 *Let* (S_L, S_R) *be an equilibrium with* $S_i = [\underline{y}_i, \bar{y}_i]$. *Then*

$$\frac{\partial}{\partial y}\gamma_i(y, [y, \bar{y}_i], S_{-i})|_{y=\underline{y}_i} \geq 0 \geq \frac{\partial}{\partial y}\gamma_i(y, [\underline{y}_i, y], S_{-i})|_{y=\bar{y}_i}.$$

Lemma 2 says that starting from an equilibrium, expanding a coalition (either to the left or to the right) must have a negative impact on the net benefit of the marginal member. This requires that, at the equilibrium, the direct negative effect of moving y on the net benefit $\gamma_i(y, S_L, S_R)$ must exceed the indirect positive effect from expanding the coalition. Otherwise, there would be a minimal expansion.

3.2 Equilibrium Existence

From Lemma 1, it follows that establishing the existence of an equilibrium requires proving that there exists a pair (S_L, S_R), with $S_i = [\underline{y}_i, \bar{y}_i]$, such that

$$\gamma_L\left(\underline{y}_L, S_L, S_R\right) = 0 = \gamma_L\left(\bar{y}_L, S_L, S_R\right), \tag{2}$$

$$\gamma_R\left(\underline{y}_R, S_L, S_R\right) = 0 = \gamma_R\left(\bar{y}_R, S_L, S_R\right), \tag{3}$$

that also satisfy the no minimal expansion condition.

We first show that an equilibrium always exists in symmetric settings.

Proposition 1 *Let $x_L = -x_R$ and the distribution ideologies, Φ, be symmetric around zero. Then a symmetric equilibrium exists.*

The key to the argument is that increases in S_i serve to raise the value for citizens of having a strategic coalition form for candidate i, thereby increasing their incentives to join that coalition (and vote). Iterating on the monotone mapping describing the marginal (indifferent) coalitional member, and applying Tarski's fixed point theorem then delivers existence.

Alternatively, we can relax the symmetry requirement on the distribution of citizens' ideologies, Φ, and candidate locations, x_L, x_R, and instead establish existence of an equilibrium when the distribution over net partisan turnout G is uniform.

Assumption 1 *The distribution G over net partisan turnout is uniform on $\left[-\frac{1}{2\alpha}, \frac{1}{2\alpha}\right]$, where $\alpha \leq 0.5$.*

Proposition 2 *Suppose that Assumption 1 holds. Then an equilibrium exists.*

As with Proposition 1, the proof of existence uses Tarski's fixed point theorem.

In most of our remaining analysis, we specialize to a uniform distribution G over net partisan turnout to simplify analysis and improve expositional clarity. The analysis is simpler because, when G is uniform, the conditions describing S_L and S_R are separable and do not depend on each other. To see this, consider the equation (2) that describes the 'responses' $\underline{y}_L$ and $\bar{y}_L$ by coalition supporters of L to S_R (i.e., (2) determines the marginal members of coalition S_L). When G is a uniform distribution with sufficient support, $P(S_L, S_R) - P(E_L, S_L)$ reduces to $\alpha(\Phi(S_L) - \Phi(E_L))$, i.e., S_R drops out of the conditions pinning down $\underline{y}_L$ and $\bar{y}_L$. This reflects the fact that with a uniform distribution, anything that shifts, say, the support of S_R on the interior does not affect the size of S_R, and hence does not affect the winning probability, which is what enters the strategic considerations of citizens on the left. It follows that the size of S_L does not vary with such shifts (and vice versa).[18]

Substituting the uniform distribution on $\left[-\frac{1}{2\alpha}, \frac{1}{2\alpha}\right]$ for $\alpha \leq 0.5$ into (2) yields that $\underline{y}_L$ and $\bar{y}_L$ must satisfy:

$$\beta(A - (x_L - y)^2) - c - (1-\beta)\alpha\left(\Phi(S_L) - \Phi(E_L)\right)\left((x_L - y)^2 - (x_R - y)^2\right) = 0, \tag{4}$$

[18] Note that this insight does not rely on the symmetry of G.

which does not depend on S_R. Analogously, if $S_R = [\underline{y}_R, \bar{y}_R] \neq \emptyset$ then $\bar{y}_R$ and $\underline{y}_R$ must satisfy:

$$\beta(A - (x_R - y)^2) - c + (1 - \beta)\alpha\left(\Phi(S_R) - \Phi(E_R)\right)\left((x_L - y)^2 - (x_R - y)^2\right) = 0, \tag{5}$$

which does not depend on S_L.

3.3 Equilibrium Characterization

In this section, we characterize the unique equilibrium for two different cases: purely cooperative motives ($\beta = 0$) and purely selfish motives ($\beta = 1$). Focusing on these extreme cases isolates the specific impact of the different motives on voters' behavior, facilitating understanding. Analyzing these cases helps shed intuition on outcomes for the intermediate case with mixed motives, i.e., $\beta \in (0, 1)$. That case proves to be mostly a weighted combination of the forces at work when $\beta = 0$ and $\beta = 1$.

A crucial result that holds in both cases is that, in equilibrium, voting coalitions S_L and S_R include a positive mass of citizens only if the cost of voting c is sufficiently small. However, both versions of the model can produce substantial turnout of non-partisan in equilibrium even when c is strictly positive. Our model thus does not suffer from the paradox of voting, a key criterion that a voting model should satisfy.

Purely Selfish Preferences

When voters have purely selfish preferences ($\beta = 1$), we obtain a closed-form solution for the equilibrium coalitions:[19]

Proposition 3 *Suppose that $\beta = 1$. There exists a unique equilibrium (S_L, S_R) defined by*

$$S_i = E_i = \left[x_i - \sqrt{A - c}, x_i + \sqrt{A - c}\right]. \tag{6}$$

This implies that, when non-empty, the expressive voter sets E_L and E_R are symmetric around x_L and x_R. This is because incentives to abstain are the same for moderates and non-moderate expressive voters. Posed differently, expressive voters on both sides of a candidate's platform can be "alienated."

Note that, E_i could include a citizen at position $y = (x_L + x_R)/2$ who is indifferent between the two candidates. This occurs when x_L and x_R are "too close" to each other. It eases presentation and analysis of the comparative statics to rule out this case in order to avoid having to augment proofs to accommodate this special case. Thus, we assume that $x_R - x_L > 2\sqrt{A - c}$. For that analysis, we will also assume that $c \leq A$, which ensures that E_L and E_R are not empty.

[19] As we show in the proof, this result does not require specific assumptions about the forms of the distributions G and Φ.

Purely Cooperative Preferences

Now consider purely cooperative preferences ($\beta = 0$). As we have seen above, the free-riding disutility of citizen y, B_y, increases with the degree of extremism $|y|$ of that citizen's ideology. It follows that the voting coalitions are unbounded in the extreme (see the proof of Proposition 4 below): if S_L is non-empty then it takes the form $S_L = (-\infty, \bar{y}_L]$, and if S_R is non-empty then it takes the form $S_R = [\underline{y}_R, \infty)$. That is, with purely cooperative preferences, non-moderate voters are never "alienated."

Under Assumption 1 that G is uniform, the expressions (4) and (5) that pin down $\bar{y}_L$ and $\underline{y}_R$ simplify to:

$$\Phi(\bar{y}_L)(x_R - x_L)(x_L + x_R - 2\bar{y}_L) - c/\alpha = 0, \quad (7)$$

$$(1 - \Phi(\underline{y}_R))(x_R - x_L)(2\underline{y}_R - x_L - x_R) - c/\alpha = 0. \quad (8)$$

From the log-concavity of the distribution Φ, we have that $\gamma_L(y, [-\infty, y], S_R)$ and $\gamma_R(y, S_L, [y, \infty])$ are single-peaked in y (see the proof of Proposition 4 below). There are thus at most two solutions to (7) and two to (8). Figure 2 illustrates two different cases for (7). In the third possible case, c is sufficiently high such that there is no solution (γ_L is always strictly below the x-axis).

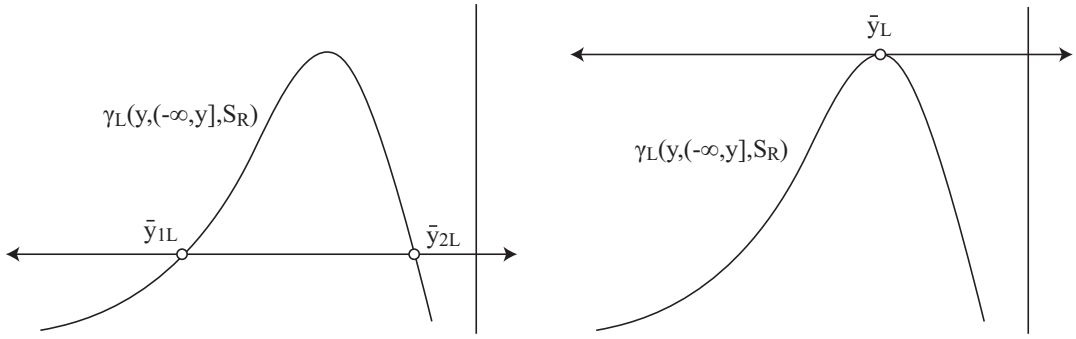


Figure 2: *Left Panel: c* sufficiently low such that there are two solutions, $\bar{y}_{L1}$ and $\bar{y}_{L2}$ *Right Panel: c* intermediate such that there is only one solution, $\bar{y}_L$.

To state our characterization result, it is useful to introduce some notation. First, with a slight abuse of notation, define $\gamma'_L(y; S_R) \equiv \frac{\partial \gamma_L(y,[-\infty,y],S_R)}{\partial y}$ and $\gamma'_R(y; S_L) \equiv \frac{\partial \gamma_R(y,S_L,[y,\infty])}{\partial y}$. Second, let $\bar{y}^*_L$ be the unique solution to $\gamma_L\left(\bar{y}^*_L, [-\infty, \bar{y}^*_L], S_R\right) = 0$ that satisfies $\gamma'_L\left(\bar{y}^*_L, S_R\right) \leq 0$, and let $\underline{y}^*_R$ be the unique solution to $\gamma_R(\underline{y}^*_R, S_L, [\underline{y}^*_R, \infty]) = 0$ that satisfies $\gamma'_R(\underline{y}^*_R; S_L) \geq 0$. Finally, define

$$\bar{c}_L = \max_{y<0.5(x_L+x_R)} \alpha\Phi(y)(x_R - x_L)(x_L + x_R - 2y), \text{ and } \bar{c}_R = \max_{y>0.5(x_L+x_R)} \alpha(1-\Phi(y))(x_R - x_L)(2y - x_L - x_R). \quad (9)$$

Proposition 4 *Suppose that Assumption 1 holds and that $\beta = 0$. There exists a unique equilibrium (S_L, S_R). S_i is non-empty if and only if $c \leq \bar{c}_i$.*

1. *If $c \leq \bar{c}_L$ then $S_L = (-\infty, \bar{y}^*_L]$, where $\bar{y}^*_L < (x_L + x_R)/2$.*

2. *If* $c \leq \bar{c}_R$ *then* $S_R = [\underline{y}^*_R, \infty)$*, where* $\underline{y}^*_R > (x_L + x_R)/2$.

Suppose that $\gamma_L(y, [-\infty, y], S_R)$ is maximized at $\bar{y}_L$. Then a solution to equation (7) exists if and only if $\gamma_L(\bar{y}_L, [-\infty, \bar{y}_L], S_R) \geq 0$, or equivalently if $c \leq \bar{c}_L$. If $c < \bar{c}_L$ then, because $\gamma_L(\bar{y}_L, [-\infty, \bar{y}_L], S_R)$ is single-peaked, equation (7) has two solutions, $\bar{y}_{L1} < \bar{y}_{L2}$. Single-peakedness of γ_L implies that $\gamma'_L(\bar{y}_{L1}; S_R) > 0 > \gamma'_L(\bar{y}_{L2}; S_R)$ (see Figure 2). In an equilibrium, by Lemma 2, it must be that $\bar{y}^*_L = \bar{y}_{L2}$—else marginal increases of the coalition would make all new members strictly better off, violating the no minimal expansion condition.

A stark feature of the equilibrium when $\beta = 0$ and G is distributed uniform is that non-moderate voters are never "alienated." This result is the extreme version of a general feature of voting equilibria, which holds even when G is not uniform: moderates are less willing to participate than non-moderates. In particular, the marginal non-moderate coalition member is located further from the candidate's platform than the marginal moderate coalition member (i.e., $x_L - \underline{y}_L > \bar{y}_L - x_L$ and $x_R - \underline{y}_R < \bar{y}_R - x_R$):

Proposition 5 *Let* $\beta < 1$*. If the coalitions* S_L *and* S_R *are non-empty, then* $\underline{y}_L$ *is further from* x_L *than* $\bar{y}_L$ *is from* x_L*, and similarly* $\bar{y}_R$ *is further from* x_R *than* $\underline{y}_R$ *is from* x_R*.*

This result simply reflects that a non-moderate coalition member is hurt by more than a moderate when her less-preferred candidate wins, as he is further away. This contrasts sharply with outcomes when citizens have purely expressive preferences.[20]

3.4 Comparative Statics of Turnout

We now vary the primitives of the electoral environment and contrast the consequences for equilibrium outcomes for the two opposite cases of purely cooperative motives ($\beta = 0$) and of purely selfish motives ($\beta = 1$). We investigate the impacts on non-partisan turnout which, for candidate i, is given by the number $\Phi(S_i)$ of non-partisans in the coalition S_i. All but the last result follow directly from Propositions 3 and 4.

While our model produces many comparative static results, for the sake of conciseness, we focus on five of them here: the impact of (i) changes in voting costs, c, (ii) changes in elite polarization, $x_R - x_L$, (iii) unilateral shifts in one candidate's platform, either x_L or x_R, (iv) shifts in the support for candidate R,

[20] In Adams and Merrill (2003) non-participation of moderates is generated by their indifference assumption, i.e., if the difference in utility between candidate L and R is less than some ϵ, then a voter abstains. In our model, both expressive and cooperative motives may drive abstention of moderates. Consider a citizen close to $(x_L + x_R)/2$. From a cooperative motive standpoint, she cares little about who wins. From an expressive motive standpoint, she may be too far from the platform of her preferred candidate to be willing to bear the cost of voting. Concerning the abstention of non-moderates, their notion of alienation reduces to our notion if (and only if) citizens only have purely selfish preferences, i.e., if $\beta = 1$.

captured by a shift in the distribution of non-partisan ideologies to the right by μ (i.e., $\Phi_\mu(x) = \Phi(x-\mu)$), and (v) changes in the weight β citizens put on their selfish vs. cooperative motives. This focus is driven by our desire to connect the predictions of our model with the empirical literature, with a focus on studies identifying a causal impact of specific features of the electoral environment on voting outcomes. The other results are relegated to Appendix 5.

We start with the impact of the cost of voting:

Result 1 *Increasing the cost of voting, c, decreases turnout for both $\beta = 0$ and $\beta = 1$.*

With purely selfish preferences ($\beta = 1$), equation (6) immediately implies that when voting costs increase, E_L and E_R shrink, and hence turnout falls. An analogous result obtains for purely cooperative preferences, $\beta = 0$. To see this, recall that the boundaries of S_L and S_R are given by the solutions to the equations $\gamma_L(\bar{y}_L, [-\infty, \bar{y}_L], S_R) = 0$ and $\gamma_R(\underline{y}_R, S_L, [\underline{y}_R, \infty]) = 0$ that satisfy $\gamma'_L(\bar{y}_L; S_R) < 0 < \gamma'_R(\underline{y}_R; S_L)$. Increasing c to c' reduces the left-hand side of those equations without affecting the right-hand side. This means that $\bar{y}_L$ and $\underline{y}_R$ must both move further away from zero, i.e., the sets S_L and S_R shrink, and hence turnout falls.

The left panel of Figure 3 illustrates the simple intuition (which does not hinge on the value of β): if the cost of voting increases, then the net benefit of the marginal coalition member jumps down. As starting from equilibrium, expanding the size of the coalition reduces the net benefit of its marginal member, to retrieve indifference of the marginal member, the coalition must shrink.

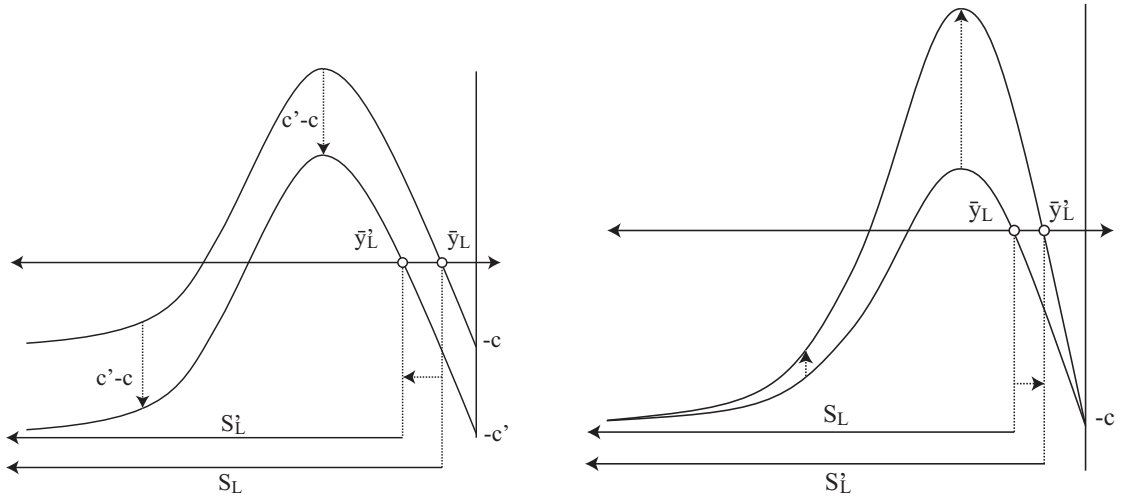


Figure 3: *Left Panel:* Increasing voting costs from c to c' reduces S_L to S'_L for $\beta = 0$. *Right Panel:* Increasing elite polarization, $x_R - x_L$ increases S_L to S'_L for $\beta = 0$.

This result is in line with empirical evidence about the positive, causal impact on turnout of features of

the electoral environment that reduce voting costs. These include early voting laws and all-mail voting laws (see, e.g., Kaplan and Yuan (2020), Thompson et al. (2020), or Yoder et al. (2021)), inclement weather on election day (Gomez et al. (2007)), proximity to polling stations (see, e.g., Cantoni (2020), Tomkins et al. (2023), and Bagwe et al. (2022)), and lax registration requirements (Braconnier et al. (2017)).[21]

Our second result focuses on the impact of increased elite polarization, i.e., increased separation between the candidates' platforms x_R and x_L.

Result 2 *Suppose that $x_L < 0 < x_R$. Increasing elite polarization $x_L - x_R$ while keeping $x_L + x_R$ fixed increases turnout when $\beta = 0$, but decreases turnout when $\beta = 1$.*

With purely cooperative preferences, shifting x_R to the right and x_L to the left by the same amount increases, e.g., $\gamma_R(y, S_L, [y, \infty])$. As a consequence, $\underline{y}_R$ must decrease to get back to equilibrium (since $\gamma'_R(\underline{y}_R; S_L) > 0$). The right panel of Figure 3 illustrates this effect. Intuitively, with purely cooperative preferences, the marginal member of a coalition has more at stake when the candidates locate further apart. In contrast, with purely selfish preferences, moving x_L to the left and x_R to the right just shifts E_L and E_R to the left and right, respectively. The single-peakedness of Φ then implies that $\Phi(E_L)$ and $\Phi(E_R)$ decrease.

The empirical literature provides mixed evidence about the impact of elite polarization on turnout (see, e.g., Béjar et al. (2020), Simas and Ozer (2021) for positive estimated effects of polarization on turnout, and Rogowski (2014), Tausanovitch and Warshaw (2018) for null or negative effects). Our model can make sense of these findings as the impact of polarization on turnout (i) varies with the motives of voters (positive if cooperative, negative if selfish), and (ii) can be small or even null if voters have mixed motives. Interpreting elite polarization as the stakes of the election for citizens, this result also speaks to the general pattern that turnout increases with the importance of the office for which candidates are running.[22]

Our third result concerns the impact of one candidate becoming more extreme:

Result 3 *Suppose that $x_L < 0 < x_R$ and that x_R becomes more extreme while x_L does not change. For $\beta = 1$, turnout for candidate L is unchanged, while turnout for candidate R decreases. For $\beta = 0$, turnout for candidate L rises, while turnout for candidate R can either increase or decrease depending on parameters.*

[21] Note that some of these changes (e.g., the introduction of mail voting) may also affect the social incentives/pressure on citizens by making the act of voting or abstaining less visible. In our model, this would mean an increase in the value of β (i.e., less weight on the cooperative motives). Such a change could have a countervailing effect on turnout (see Proposition 6 below), as discussed and tested in Funk (2010).

[22] For instance, in the US, turnout is systematically higher in Presidential than in Senatorial or House elections (this is true both comparing across races for any given on-cycle election due to voter roll-off, and comparing across races in different electoral cycles, with turnout dropping substantially in mid-term elections).

The results for purely selfish preferences ($\beta = 1$) follow immediately from the characterization of E_L and E_R (the former is independent of x_L and the latter is centered around x_R) as well as the fact that the distribution Φ is single peaked.

For purely cooperative preferences ($\beta = 0$), we need to understand the impact of the increase of x_R on $\bar{y}_L$ and $\underline{y}_R$. Note that $\gamma_L(y, [-\infty, y], S_R)$ is increasing in x_R. Hence, the same argument as for the previous result yields that the relevant solution of the equation $\gamma_L(\bar{y}_L, [-\infty, \bar{y}_L], S_R) = 0$ must increase. Intuitively, candidate R's victory is a worse outcome for leftist voters when x_R is more extreme, thus boosting turnout. In contrast, the impact on $\underline{y}_R$ is ambiguous. Recall that $\gamma_R(y, S_L, [y, \infty]) \equiv \alpha(1-\Phi(y))(x_R - x_L)(2y - x_L - x_R) - c$. It is immediate that $\partial\gamma_R/\partial x_R$ is positive if $y > x_R$ and negative otherwise. Intuitively, candidate R's victory is a worse outcome for the marginal coalition member $\underline{y}_R$ if a shift of x_R to the right means moving away from $\underline{y}_R$, but it is a better outcome for that voter if it means moving closer to $\underline{y}_R$.[23]

Under the assumption that, on average, candidates are more extreme than the most moderate citizen voting for them, these results align with the findings in the literature.[24] In particular, Hall and Thompson (2018) and Hall and Thompson (2025) use a regression discontinuity design on the result of the primaries in US House races to show that the quasi-random nomination of a more extreme candidate (instead of a more moderate one) reduces the turnout of supporters and increase the turnout of opponents in general elections.

Now consider the effect of moving the distribution of non-partisans by μ to the right to $\Phi_\mu(x) = \Phi(x - \mu)$.

Result 4 *Suppose the distribution of non-partisan ideologies is shifted to the right. For $\beta = 0$, turnout for L falls and turnout for R rises. For $\beta = 1$, there is no effect on E_L and E_R, and hence the impact on turnout for L and R depends on the distribution of voter ideal points, Φ, and the positions of the candidates, x_L and x_R.*

For purely cooperative preferences ($\beta = 0$), inspection reveals that increasing μ is equivalent to reducing $x_L + x_R$, leaving $x_R - x_L$ unaffected, thereby decreasing $\gamma_L(\bar{y}_L, [-\infty, \bar{y}_L], S_R)$ and increasing $\gamma_R(\underline{y}_R, S_L, [\underline{y}_R, \infty])$. Thus, both $\bar{y}_L$ and $\underline{y}_R$ must decrease. For purely selfish preferences ($\beta = 1$), it is clear from the inspection

[23] A naïve intuition might suggest that x_R is always larger than $\underline{y}_R$ because a voter whose ideal point is given by the policy x_R should want to be part of a coalition supporting R. However, when candidate policies are not very differentiated, i.e., x_L and x_R are quite close, citizen $y = x_R$ may not care enough about who wins to benefit much from the voting efforts of others. In contrast, given the quadratic loss assumption, citizens who are sufficiently extreme care more about who wins. If there are sufficiently many such extreme voters, i.e., if the standard deviation of Φ is sufficiently large, then these extreme citizens are incentivized to vote for R.

[24] While there is a relatively large empirical literature exploring whether voters punish candidates who take more extreme ideological positions, the so-called "electoral penalty" (see Bailey and Reese (2025), Bonica et al. (2025) and references therein), most of these studies do not directly speak to our result because they focus on the impact of candidate platforms on their vote shares. As such, they not only capture the impact of extreme platforms on turnout (which is directly related to the prediction of our model), but also the behavior of citizens when at the poll booth (which our model has nothing to say about). For instance, both Bailey and Reese (2025) and Bonica et al. (2025) focus exclusively on the latter, by exploring the behavior of voters across electoral races on a given ballot. Their results are thus not directly related to our result about the impact of platform extremeness on turnout.

of equation (6) that the shifts does not affect E_L and E_R. This is because selfish citizens do not weigh win probabilities in their decision-making. In terms of turnout, the shift (generically) leads to a rise in the turnout for one candidate and a fall for the other, but which candidate gains support and which one loses depends on the distribution of voter ideal points and the platforms of the candidates.

To see the intuition of the result for cooperative preferences, consider the turnout for candidate L. The move of x_R to the left means that the marginal coalition member for L (before the move) finds a victory of candidate R not as bad an outcome. This reduces her incentives to join the coalition. The same-sized move of x_L to the left may mean that this marginal coalition voter finds a victory of candidate L a more desirable outcome (if $x_L > \bar{y}_L$), which would increase her incentives to join the coalition. However, even when the two effects push in opposite directions, the impact of the change in x_R dominates due to the quadratic loss assumption (as candidate R is further away from the marginal L voter than candidate L is), so the overall effect on the incentives to join the coalition is negative. The same logic yields that turnout for R falls.

These results relate to the literature on the effect on turnout of a change in the closeness of the electoral support for the two candidates. Using different empirical strategies, this literature finds effects of closeness ranging from non-significant to positive but small (Moskowitz et al. (2019) and Ainsworth et al. (2024) on US data, and Bursztyn et al. (2024) on Swiss data). Our model proposes an explanation for these findings: the effect of increasing closeness on overall turnout is small simply because it affects turnout for L and R in opposition directions. This stands in stark contrast with predictions of other voting models, which typically predict a positive impact of closeness on the turnout of supporters of both candidates, and hence on overall turnout. Our results also call for a test of heterogeneous effects of closeness on the turnout of supporters of the frontrunner and those of the runner-up.

While this section focuses on the comparative statics of turnout for the benchmark cases of purely selfish ($\beta = 1$) and purely cooperative ($\beta = 0$) preferences, we can also investigate how the value of β affects turnout for values of β in the vicinity of these benchmark settings.

Proposition 6 *Let Assumption 1 hold and suppose that $c < \bar{c}_i$. For β close to zero, there is an $\bar{A}$ such that turnout increases in β if $\bar{A} > A$, but decreases otherwise. For β close to* 1, *turnout increases in β.*

For β close to 1, the intuition is simple. As citizens have (almost) no cooperative motives, turnout is (essentially) fully driven by the selfish motive ($S_i = E_i$, $\forall i \in \{L, R\}$). Increasing β increases the expressive payoff from voting relative to the cost of voting. As a result, the set of citizens who vote for expressive reasons expands.

For β close to 0, the effect of an increase in β depends on the difference between the expressive payoff of voting and the free-riding disutility for the marginal coalition member, as well as those with similar ideologies. For these (near) marginal citizens, if the expressive payoff of voting is relatively larger (e.g.,

when A is large enough), then increasing the weight citizens assign to this motive when deciding whether to vote raises the net benefit of voting, say, $\gamma_L(y, [-\infty, y], S_R)$. The coalition $S_L = (-\infty, \bar{y}_L]$ thus has to expand to retrieve equilibrium, characterized by the fixed point, $\gamma_L\left(\bar{y}_L, (-\infty, \bar{y}_L], S'_R\right) = 0$. By contrast, if the free-riding disutility is relatively larger (e.g., when A is small enough), then increasing the weight citizens assign to this motive when deciding whether to vote reduces the net benefit of voting. The coalition S_L thus has to shrink in order to return to equilibrium.

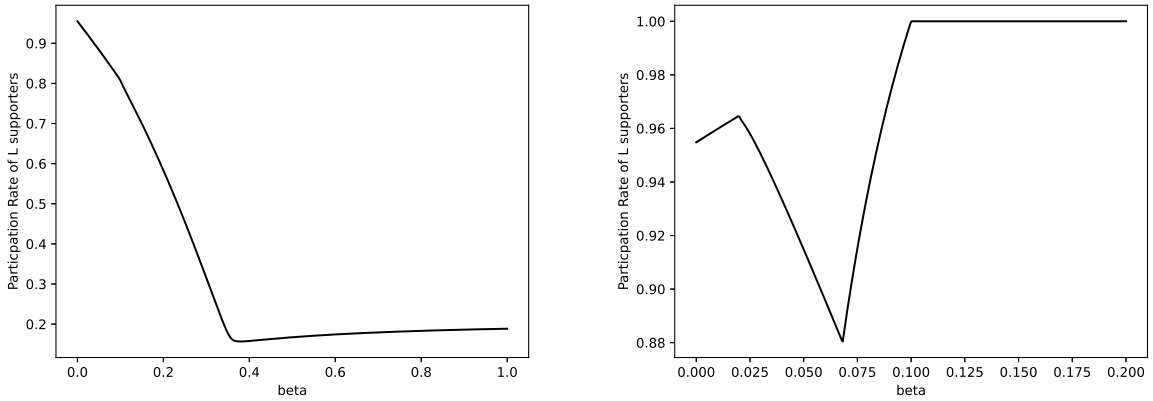


Figure 4: Participation rates for candidate L as a function of β for $A = 1$ (left panel) and $A = 5$ (right panel). Φ is normal and G uniform, with standard deviations of 0.5 and 0.8, respectively. Other Parameters: $x_R = -x_L = 2$, $c = 0.1$.

A direct implication of Proposition 6 is that as long as the purely expressive benefit A of voting is not quite large, equilibrium turnout is not monotone in β. As the left panel of Figure 4 illustrates, when A is sufficiently small, starting from $\beta = 0$, equilibrium turnout first decreases. It must then start to increase for values of β sufficiently large as we know that equilibrium turnout is increasing in β when β is close to 1. As the right panel of Figure 4 shows, this non-monotonicity does not require A to be small. For the case depicted, A is large enough that equilibrium turnout is increasing in β when β is close to zero. As explained above, this is because citizens put more weight on the expressive payoff of voting, which exceeds the free-riding disutility. But, the figure also shows that with further increases in β, equilibrium turnout starts to go down before going up again. The first inflection stems from the indirect effect that β has on the disutility of free-riding via its effect on the set of expressive voters supporting candidate L, E_L. In particular, E_L is empty for sufficiently low values of β, but is non-empty and increasing in β for higher values of β. When that set expands, the disutility of free-riding decreases because it reduces the impact of coalition S_L on the electoral outcome (i.e., $\Phi(S_L) - \Phi(E_L)$ decreases). This reduces the incentives of the marginal member of S_L to vote. The second inflection stems from the end of "cooperative voting", i.e., it is the value of β at which $S_L = E_L$. The positive impact of β on equilibrium turnout then follows from the fact that E_L is increasing in β.

Under the assumption that citizens in societies with higher social capital put more weight on cooperative motives and less on selfish motives (i.e., lower β), our non-monotonicity result can help reconcile the findings of the empirical literature exploring the link between social capital and turnout. Most of that literature finds evidence of a positive correlation between social capital and turnout (see the discussion in Atkinson and Fowler (2012)). However, the few papers that study the causal impact of social capital on turnout reach mixed conclusions: Condon (2009) and Atkinson and Fowler (2012) find a negative impact, whereas Gerber et al. (2016) finds a positive impact. Our non-monotonicity result also casts doubts on the wisdom of the common practice of proxying social capital with measures of turnout (see Durante et al. (2025)).

3.5 Comparative Statics of Turnout Rate: Bandwagon and Underdog Effects

Previous theoretical research on costly voting predicts the existence of an underdog effect, i.e., supporters of the runner-up turn out at a higher rate than supporters of the front-runner. For instance, using the pivotal voter model, Palfrey and Rosenthal (1985) assume that a particular citizen is a candidate L supporters with probability p and a R supporter with probability $1 - p$. Voters have a cost drawn from some distribution F, and the Nash equilibrium is determined by cost cutoffs c_A and c_B for A and B supporters, respectively. If $p < 0.5$, then $c_A > c_B$, i.e., participation rates are higher among supporters of the runner-up. For group-based voting models, Feddersen and Sandroni (2006) also predict an underdog effect.

This prediction has been tested in the laboratory and most studies (Duffy and Tavits (2008); Großer and Schram (2010); Kartal (2015); Agranov et al. (2018)) find the opposite result: supporters of the front-runner turn out at a *higher rate* than supporters of the runner-up—the so-called *bandwagon effect*.[25] Agranov et al. (2018) argue that a modification of the pivotal voter model in which voters like to vote for the winner, as in Callander (2008), organizes their data well, but that neither other variants of the pivotal voter model (e.g., Grillo (2017)) nor the ethical voter model (Feddersen and Sandroni (2006)) do. As we show below, our model provides an alternative explanation for these findings of bandwagon effects: cooperatively-motivated voters are subject to a discouragement effect when they anticipate having fewer co-supporters in the electorate as the disutility of free riding then decreases.[26]

[25]The only exception we are aware of is Levine and Palfrey (2005), which find evidence of an underdog effect in a laboratory experiment. Note however that, as discussed in Agranov et al. (2018), the effect they observe is quite small.

[26]Testing the underdog effect using observational data is challenging as it requires to measure the turnout *rate* of supporters of the front-runner and the runner-up. The only relevant analysis using observational data that we know of is in the Appendix of Bouton et al. (2023). Using a differences-in-differences approach relying on changes in competitiveness at the precinct level following the 2020 redistricting cycle, they find fairly precise null effects of partisan alignment (i.e., underdog or bandwagon effect) on turnout. Our model provides an explanation for these findings: the bandwagon effect stemming from social incentives is compensated by the underdog effect that may stem from selfish incentives.

We define turnout rates r_i for candidates $i = L, R$ as follows:

$$r_L = \frac{\Phi(S_L)}{\Phi((x_L + x_R)/2)} \text{ and } r_R = \frac{\Phi(S_R)}{1 - \Phi((x_L + x_R)/2)}. \tag{10}$$

The denominators are the number of citizens whose bliss point is closer to a particular candidate, i.e., the number of citizens who expressively prefer that candidate. This is the closest analogue in our model to the exogenous probability p that a particular voter is an L supporter.

To model changes in the number of potential L and R supporters we shift the distribution Φ of citizens as in Result 4. From this result, we know that with purely cooperative preferences ($\beta = 0$), a shift of the citizen distribution that is favorable to candidate R reduces the turnout for L, $\Phi(S_L)$, which has a negative effect on the turnout rate. However, the shift also reduces the number of potential L supporters, $\Phi((x_L + x_R)/2)$, which ceteris paribus would increase the turnout rate of L supporters. Proposition 7 below shows that, starting from a symmetric situation, the first effect always dominates: turnout for the disadvantaged candidate falls, while it rises for the advantaged candidate. The proposition also contrasts this outcome with the case of purely selfish citizens. From Result 4, we know that following a shift in the distribution Φ in favor of candidate R, the turnout for L, $\Phi(E_L)$, goes down. As the shift also reduces the number of potential supporters for candidate L, the overall effect on the turnout rate r_L depends on the extent of mass polarization, as measured by the standard deviation σ of Φ. The same holds for the effect on the turnout rate r_R.

Proposition 7 *Suppose that Assumption 1 holds, that Φ is symmetric around 0, and that $x_R = -x_L$. Further, suppose that $c_i = c < \bar{c}_i$, $i = L, R$ so that a non-trivial equilibrium exists. Let $\Phi_\mu(x) = \Phi(x - \mu)$, and increase μ marginally above 0, thereby providing an advantage to candidate R. Then:*

1. *For $\beta = 0$, if $c < \bar{c}$, the turnout rate r_L falls and the turnout rate r_R rises.*
2. *For $\beta = 1$, if the standard deviation of Φ is sufficiently large, then the turnout rate r_L rises and the turnout rate r_R falls.*

Thus, if citizens have sufficiently cooperative preferences, Proposition 7 implies that candidates who are believed to have more supporters gain further from a bandwagon effect that increases participation by citizens who would have otherwise sat out the election. Similarly, if a candidate falls behind, this effect is reinforced, as citizens who were previously willing to incur costs of voting, are now discouraged and stay home.

4 Endogenous Platforms

In this section, we show how the underlying motives of voters affect candidates' choices of platforms. We endogenize the platforms of the candidates in a classical Calvert-Wittman style model of electoral competi-

tion between two policy-motivated candidates, contrasting outcomes when citizens hold purely cooperative ($\beta = 0$) and purely selfish ($\beta = 1$) preferences. We maintain Assumption 1 on the distribution of partisan voters, and focus on a symmetric setting (i.e., Φ symmetric around 0).

Candidates have quadratic preferences, with ideal points denoted by θ_i, $i = L, R$, where $\theta_L = -\theta_R$. Let $Q(x_L, x_R)$ be the probability that candidate L wins given platform choices x_L and x_R by candidates L and R. Candidate R solves

$$\max_{x_R} -Q(x_L, x_R)(x_L - \theta_R)^2 - (1 - Q(x_L, x_R))(x_R - \theta_R)^2.$$

For purely cooperative preferences, we have shown that

$$Q(x_L, x_R) = G(\Phi(\bar{y}_L) - (1 - \Phi(\underline{y}_R))),$$

where $\bar{y}_L$ and $\underline{y}_R$ solve (7) and (8), respectively. For purely selfish preferences, we have shown that

$$Q(x_L, x_R) = G(\Phi(E_L) - \Phi(E_R)),$$

where E_i is given by (6).

4.1 Equilibrium Characterization

We first characterize how the underlying motives of voters shape the choice of platforms by the candidates.

Proposition 8 *There exists $\bar{c} > 0$ such that if voting costs $c < \bar{c}$ then candidates choose more extreme platforms when $\beta = 1$ than when $\beta = 0$.*

To understand why candidates adopt more extreme platforms when citizens have purely selfish preferences, note that the effect of moving a candidate's platform closer to his bliss point can be decomposed into two parts. First, there is a negative impact on the probability of winning the election. Second, there is a positive impact on the candidate's utility of winning with a platform closer to his bliss point. The second effect does not vary with the behavior of voters, as it only depends on the utility of the candidates. Comparisons of the first effect for the cases with purely cooperative and purely selfish preferences thus drive the result.

The key is that, for the purely selfish case, there are both moderates and non-moderate citizens who are alienated. A move of a candidate's platform closer to her bliss point then alienates additional moderate citizens but also convinces some non-moderate citizens to vote instead of abstain. These opposite effects limit the impact of the change in platform on the probability that the candidate wins, which frees candidates to move closer to their bliss points.[27] By contrast, with purely cooperative preferences, only moderate citizens

[27]For the case of uniform Φ, the two effects exactly compensate and there is thus no impact of the platform change on the probability of winning. Candidates then always adopt their ideal points as policies.

are alienated, while all sufficiently non-moderate citizens join the voting coalition. Moving a candidate's platform closer to her bliss point then alienates additional moderate citizens, but does not help convince any non-moderate ones to join the coalition. The impact on the probability of winning is thus larger, which limits the incentives of the candidates to move closer to their bliss point, reducing polarization.

One direct implication of Proposition 8 is that, if citizens' behavior is (at least partially) driven by cooperative incentives, then there will be partial but not full convergence of candidates' platforms. This is in line with empirical evidence, which shows that (i) candidates' platforms rarely fully converge (i.e., there is polarization) and that (ii) candidates are reactive to voters' ideologies in the sense that their platforms move with the position of the expected median (Ansolabehere et al. (2001), Bonica (2014), Di Tella et al. (2023)).

4.2 Comparative Statics of Polarization

We are now in position to explore how features of the electoral environment affect political polarization. In this section, we investigate the impact of voting costs on elite polarization. We obtain findings that stand in stark contrast with the predictions of existing models of electoral competition, illustrating how including strategic voters in the analysis of candidates' behavior can dramatically alter conclusions.

The following Proposition shows that a change in voting costs affects polarization differently according to whether cooperative or selfish incentives drive the behavior of voters.

Proposition 9 *There exists $\bar{c} > 0$ such that if voting costs $c < \bar{c}$ then an increase in voting costs decreases elite polarization ($x_R - x_L$) if $\beta = 0$ while the reverse is true if $\beta = 1$.*

To see the impacts of increased voting costs, first note that with purely selfish preferences, increasing c reduces the size of the expressive voting coalitions E_L and E_R. This, in turn, reduces $\frac{\partial P(x_L,x_R)}{\partial x_R}$, i.e., choosing more extreme platforms reduces the winning probability by less than before the increase in c. This is because the sensitivity of the probability of winning depends on the difference between the loss in moderate votes and the gain in non-moderate votes following a move to the extreme. As Φ is single-peaked, this difference is increasing in the size of the coalition. Thus, following an increase in c, candidates have stronger incentives to adopt more extreme platforms because the cost of doing so in terms of votes (and hence probability of winning) is lower.

Now consider increasing voting costs in the purely cooperative case. To ease exposition, consider a case for which, before the change in c, $x_L \in S_L$ and $x_R \in S_R$. As we have seen above, increasing c reduces both S_L and S_R. All else equal, this implies that the marginal opponent is further away from, say, x_R, and the marginal supporter is closer. With quadratic loss preferences, this implies that $\bar{y}_L$ becomes more sensitive to a shift in x_R (i.e., $\frac{\partial \bar{y}_L}{\partial x_R}$ increases) and that $\underline{y}_R$ becomes less sensitive (i.e., $\frac{\partial \underline{y}_R}{\partial x_R}$ decreases). Importantly, the

change is larger for the former than the latter. Hence, overall, we have that the change in the probability that R wins following an increase in x_R increases (i.e., $\frac{\partial \bar{y}_L}{\partial x_R} + \frac{\partial y_R}{\partial x_R}$ increases). This means that for a candidate, moving her platform outward, closer to her bliss point, becomes more costly in terms of probability of victory. This gives candidates incentives to adopt more moderate platforms, i.e., to polarize less.

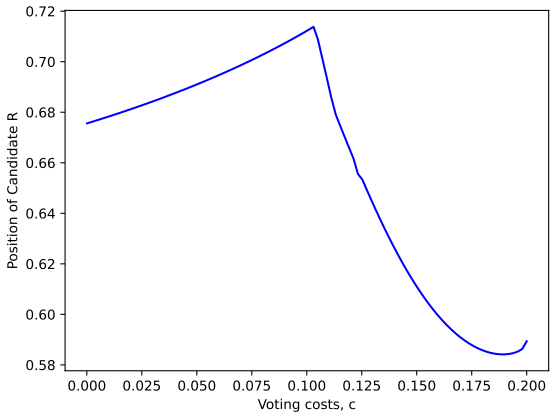


Figure 5: Candidate R's position for $\beta = 0.2$, when $\theta_R = -\theta_L = 1$, G and Φ are $N(0, 1)$ distributed. Political polarization is $x_R - x_L = 2x_R$.

Proposition 9 also has interesting policy implications which contrast with common beliefs in the literature. In particular, Oprea et al. (2024) use models of electoral competition without strategic voters to argue in favor of compulsory voting (a policy that is equivalent to a sharp reduction in voting costs) is that it should help decrease elite polarization. By contrast, our model predicts that compulsory voting could actually *increase* polarization if voters have sufficiently cooperative preferences. We are not aware of any existing empirical evidence that could directly inform this theoretical debate. Such work could either explore the impact of the introduction of compulsory voting on elite polarization, or other changes in the electoral environment that influence the voting costs of (most) voters.

The basic insights of Proposition 9 extend to hold for intermediate β. Figure 5 illustrates how candidate R's position varies with voting costs for $\beta = 0.2$. Note that as we increase c from low levels, policies become more extreme. The reason is that for small values of c only expressive voting occurs (and for $c = 0$ everyone votes). Thus, it is as if we are in the case $\beta = 1$ described above. Raising c lowers the number of expressive voters, and thereby increases polarization. Once c hits a critical value, strategic voting starts, and we are in a similar situation as for the case of $\beta = 0$ described in the Proposition. That is, competition for voters in the center induces candidates to moderate, until we reach a point where few citizens vote, and partisans matter more for the electoral outcome.

5 Conclusion

This paper addresses the longstanding challenge of modeling voter behavior by proposing a novel model that captures the aversion of citizens to free riding on the costly voting efforts of others. In contrast to existing group-based models, the set of voters who support a particular candidate is determined endogenously, using a standard Nash equilibrium concept. By endogenizing the formation of voting coalitions, we offer a more realistic and empirically consistent framework for understanding voter turnout and decision-making. Our model not only overcomes the limitations of existing models but also provides a clear mechanism for how and why citizens coordinate their voting behavior in response to the characteristics of the electoral environment. Further, the model is sufficiently tractable to allow us to endogenize candidate platform choices, and thus show how platforms and elite polarization hinge on the nature of voter preferences and voting costs.

References

Abramson, P. R., J. H. Aldrich, A. Diskin, A. M. Houck, R. Levine, T. J. Scotto, and D. B. Sparks (2018). The effect of national and constituency expectations on tactical voting in the british general election of 2010. In *The many faces of strategic voting*, pp. 28. University of Michigan Press Ann Arbor, MI.

Adams, J. and S. Merrill (2003). Voter turnout and candidate strategies in American elections. *Journal of Politics 65*, 161–189.

Agranov, M., J. K. Goeree, J. Romero, and L. Yariv (2018). What makes voters turn out: The effects of polls and beliefs. *Journal of the European Economic Association 16*(3), 825–856.

Ainsworth, R., E. G. Munoz, A. M. Gomez, et al. (2024). District competitiveness increases voter turnout: evidence from repeated redistricting in north carolina. *Quarterly Journal of Political Science 19*(4), 387–432.

Ali, S. N. and C. Lin (2013, May). Why people vote: Ethical motives and social incentives. *American Economic Journal: Microeconomics 5*(2), 73–98.

Ambrus, A. and P. Pathak (2011). Cooperation over finite horizons: A theory and experiments. *Journal of Public Economics 95*, 500–512.

Andreoni, J. and D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*, 1607–1636.

Andreoni, J. and J. Miller (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica 70*, 737–753.

Ansolabehere, S., J. M. Snyder Jr, and C. Stewart III (2001). Candidate positioning in us house elections. *American Journal of Political Science*, 136–159.

Atkinson, M. D. and A. Fowler (2012). The effect of social capital on voter turnout: Evidence from saints day fiestas in mexico. *University of California, Los Angeles and Harvard University*.

Bagwe, G., J. Margitic, and A. Stashko (2022). Polling place location and the costs of voting. Technical report, Working Paper.

Bailey, M. and B. Reese (2025). Ideological moderation and success in us elections, 2020-2022. *Available at SSRN 5169898*.

Battigalli, P. and M. Dufwenberg (2007). Guilt in games. *American Economic Review 97*, 170–176.

Battigalli, P. and M. Dufwenberg (2009). Dynamic psychological games. *Journal of Economic Theory 144*, 1–35.

Béjar, S., J. A. Moraes, and S. López-Cariboni (2020). Elite polarization and voting turnout in latin america, 1993–2010. *Journal of Elections, Public Opinion and Parties 30*(1), 1–21.

Belli, R., M. Traugott, M. Young, and K. McGonagle (1999). Reducing vote overreporting in surveys: Social desirability, memory failure, and source monitoring. *Public Opinion Quarterly 63*, 90–108.

Bernhardt, D., S. Krasa, and F. Squintani (2024). Political competition and strategic voting in multi-candidate elections. Working Paper, University of Warwick.

Blais, A. (2000). *To vote or not to vote: The merits and limits of rational choice theory*. University of Pittsburgh Press.

Bolton, G. and A. Ockenfels (2000). Erc: a theory of equity, reciprocity, and competition. *American Economic Review 90*, 166–193.

Bond, R. M., C. J. Fariss, J. J. Jones, A. D. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler (2012). A 61-million-person experiment in social influence and political mobilization. *Nature 489*(7415), 295–298.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science 58*(2), 367–386.

Bonica, A., K. Rhee, and N. Studen (2025). The electoral consequences of ideological persuasion: Evidence from a within-precinct analysis of us elections. *Available at SSRN*.

Bouton, L., M. Castanheira, and L.-S. Aniol (2015). Multicandidate elections: Aggregate uncertainty in the laboratory. *Games and Economic Behavior 101*, 132–150.

Bouton, L., G. Genicot, M. Castanheira, and A. L. Stashko (2023). Pack-crack-pack: gerrymandering with differential turnout. Technical report, National Bureau of Economic Research.

Braconnier, C., J.-Y. Dormagen, and V. Pons (2017). Voter registration costs and disenfranchisement: experimental evidence from france. *American Political Science Review 111*(3), 584–604.

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

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*(2), 876–914.

Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *American Economic Review 96*, 1652–1678.

Callander, S. (2008). Majority rule when voters like to win. *Games and Economic Behavior 64*(2), 393–420. Special Issue in Honor of Michael B. Maschler.

Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science 29*, 69–95.

Cantoni, E. (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics 12*(1), 61–85.

Castanheira, M. (2003). Victory margins and the paradox of voting. *European Journal of Political Economy 19*(4), 817–841.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics 117*, 817–869.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics 14*, 47–83.

Coate, S. and M. Conlin (2004). A group rule-utilitarian approach to voter turnout: Theory and evidence. *American Economic Review 94*(5), 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*(3-4), 582–596.

Condon, M. (2009). The effect of social capital on voter turnout: A field experiment in two southwestern cities. *Working Paper*.

Cubitt, R. P., M. Drouvelis, S. Gächter, and R. Kabalin (2011). Moral judgments in social dilemmas: How bad is free riding? *Journal of Public Economics 95*(3-4), 253–264.

Daoust, J.-F. (2018). Support for minority government and strategic voting. In J. H. Aldrich, A. Blais, and L. B. Stephenson (Eds.), *The many faces of strategic voting*, Chapter 4, pp. 75–88. University of Michigan Press Ann Arbor, MI.

DellaVigna, S., J. A. List, U. Malmendier, and G. Rao (2016). Voting to tell others. *Review of Economic Studies 84*(1), 143–181.

Di Tella, R., R. Kotti, C. Le Pennec, and V. Pons (2023). Keep your enemies closer: strategic platform adjustments during us and french elections. Technical report, National Bureau of Economic Research Cambridge, MA.

Doherty, D., C. Dowling, A. Gerber, and G. Huber (2017). Are voting norms conditional? how electoral context and peer behavior shape the social returns to voting. *Journal of Politics 79*, 1095–1100.

Downs, A. (1957). *An economic theory of democracy*. Harper.

Duffy, J. and M. Tavits (2008). Beliefs and voting decisions: a test of the pivotal voter model. *American Journal of Political Science 52*(3), 603–618.

Durante, R., N. Mastrorocco, L. Minale, and J. M. Snyder (2025). Unpacking social capital. *The Economic Journal 135*(667), 773–807.

Esponda, I. and E. Vespa (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Review 6*(4), 180–202.

Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and Economic Behavior 54*, 293–315.

Feddersen, T., S. Gailmard, and A. Sandroni (2009). Moral bias in large elections: Theory and experimental evidence. *American Political Science Review 103*, 175–192.

Feddersen, T. and A. Sandroni (2006). A theory of participation in elections. *American Economic Review 96*(4), 1271–1282.

Fehr, E. and S. Gächter (2002). Altruistic punishment in humans. *Nature 415*(6868), 137–140.

Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review 90*, 980–994.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of economics 114*(3), 817–868.

Fehr, E. and I. Schurtenberger (2018). Normative foundations of human cooperation. *Nature human behaviour 2*, 458–468.

Fischbacher, U. and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review 100*(1), 541–556.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics letters 71*(3), 397–404.

Forsythe, R., T. Rietz, R. Myerson, and R. Weber (1996). An experimental study of voting rules and polls in three-candidate elections. *International Journal of Game Theory 25*, 355–383.

Fowler, J. (2006). Altruism and turnout. *Journal of Politics 68*, 674–683.

Fujiwara, T. (2011). A regression discontinuity test of strategic voting and duverger's law. *Quarterly Journal of Political Science 6*, 197–233.

Funk, P. (2010). Social incentives and voter turnout: evidence from the swiss mail ballot system. *Journal of the European economic association 8*(5), 1077–1103.

Gachter, S. and B. Herrmann (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society 364*, 791–806.

Gerber, A. and T. Rogers (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *Journal of Politics 71*, 178–191.

Gerber, A. S., D. P. Green, and C. W. Larimer (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American political Science review 102*(1), 33–48.

Gerber, A. S., J. Gruber, and D. M. Hungerman (2016). Does church attendance cause people to vote? using blue laws' repeal to estimate the effect of religiosity on voter turnout. *British Journal of Political Science 46*(3), 481–500.

Gerber, A. S., G. A. Huber, D. Doherty, and C. M. Dowling (2016). Why people vote: estimating the social returns to voting. *British Journal of Political Science 46*(2), 241–264.

Gerber, A. S., G. A. Huber, A. H. Fang, and A. Gooch (2017). The generalizability of social pressure effects on turnout across high-salience electoral contexts: Field experimental evidence from 1.96 million citizens in 17 states. *American Politics Research 45*(4), 533–559.

Gomez, B. T., T. G. Hansford, and G. A. Krause (2007). The republicans should pray for rain: Weather, turnout, and voting in us presidential elections. *Journal of Politics 69*(3), 649–663.

Good, I. J. and L. S. Mayer (1975). Estimating the efficacy of a vote. *Behavioral Science 20*(1), 25–33.

Grillo, A. (2017). Risk aversion and bandwagon effect in the pivotal voter model. *Public Choice 172*(3), 465–482.

Großer, J. and A. Schram (2010). Public opinion polls, voter turnout, and welfare: An experimental study. *American Journal of Political Science 54*(3), 700–717.

Hall, A. B. and D. M. Thompson (2018). Who punishes extremist nominees? candidate ideology and turning out the base in us elections. *American Political Science Review 112*(3), 509–524.

Hall, A. B. and D. M. Thompson (2025). Revising 'who punishes extremist nominees? candidate ideology and turning out the base in us elections'.

Harsanyi, J. (1977). Morality and the theory of rational behavior. *Social Research 44*(4), 623–655.

Herrera, H. and C. Martinelli (2006). Group formation and voter participation. *Theoretical Economics 1*, 461–481.

Jankowski, R. (2007). Altruism and the decision to vote. *Rationality and Society 19*, 5–34.

Kaplan, E. and H. Yuan (2020, January). Early voting laws, voter turnout, and partisan vote composition: Evidence from ohio. *American Economic Journal: Applied Economics 12*(1), 32–60.

Kartal, M. (2015). Laboratory elections with endogenous turnout: proportional representation versus majoritarian rule. *Experimental Economics 18*, 366–384.

Kelley, H. H. and A. J. Stahelski (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of personality and social psychology 16*(1), 66.

Kimbrough, E. O. and A. Vostroknutov (2016). Norms make preferences social. *Journal of the European Economic Association 14*, 608–638.

Kiss, H., A. Rosa-Garcia, and V. Zhukova (2020). Conditional cooperation in group contests. *PLoS ONE 15*, 1–13.

Knack, S. (1992). Civic norms, social sanctions, and voter turnout. *Rationality and Society 4*, 133–156.

Kocher, M., T. Cherry, S. Kroll, R. Netzer, and M. Sutter (2008). Conditional cooperation on three continents. *Economics Letters 101*, 175–178.

Ledyard, J. (1984). Pure theory of two-candidate competition. *Public Choice 44*, 7–41.

Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics 1*, 593–622.

Levine, D. and A. Mattozzi (2020). Voter turnout with peer punishment. *American Economic Review 110*(10), 3298–3314.

Levine, D. and T. Palfrey (2005). The paradox of voter participation? A laboratory study. Princeton.

Malmendier, U., V. L. te Velde, and R. A. Weber (2014). Rethinking reciprocity. *Annu. Rev. Econ. 6*(1), 849–874.

Markus, G. (1988). The impact of personal and national economic conditions on the presidential vote: A pooled cross-sectional analysis. *American Journal of Political Science 32*, 137–154.

Merlo, A. and T. Palfrey (2018). External validation of voter turnout models by concealed parameter recovery. *Public Choice 176*, 297–314.

Morton, R. (1991). Groups in rational turnout models. *American Journal of Political Science 35*(3), 758–776.

Moskowitz, D. J., B. Schneer, et al. (2019). Reevaluating competition and turnout in us house elections. *Quarterly Journal of Political Science 14*(2), 191–223.

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

Opp, K.-D. (2001). Why do people vote? the cognitive-illusion proposition and its test. *Kyklos 54*, 355–378.

Oprea, A., L. Martin, and G. H. Brennan (2024). Moving toward the median: Compulsory voting and political polarization. *American Political Science Review*, 1–15.

Palfrey, T. R. and H. Rosenthal (1983). A strategic calculus of voting. *Public Choice 41*, 7–53.

Palfrey, T. R. and H. Rosenthal (1985). Voter participation and strategic uncertainty. *American Political Science Review 79*, 62–78.

Pons, V. and C. Tricaud (2018). Expressive voting and its cost: Evidence from runoffs with two or three candidates. *Econometrica 86*(5), 1621–1649.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review 83*, 1281–1302.

Riker, W. H. and P. C. Ordeshook (1968). A theory of the calculus of voting. *American Political Science Review 62*(1), 25–42.

Roemer, J. E. (2010). Kantian equilibrium. *Scandinavian Journal of Economics 112*(1), 1–24.

Roemer, J. E. (2015). Kantian optimization: A microfoundation for cooperation. *Journal of Public Economics 127*, 45–57.

Rogowski, J. C. (2014). Electoral choice, ideological conflict, and political participation. *American Journal of Political Science 58*(2), 479–494.

Rotemberg, J. (2008). Minimally acceptable altruism and the ultimatum game. *Journal of Economic Behavior and Organization 66*, 457–476.

Shachar, R. and B. Nalebuff (1999). Follow the leader: Theory and evidence on political participation. *American Economic Review 89*(3), 525–547.

Silver, B., B. Anderson, and P. Abramson (1986). Who overreports voting? *American Political Science Review 80*, 613–624.

Simas, E. N. and A. L. Ozer (2021). Polarization, candidate positioning, and political participation in the us. *Electoral Studies 73*, 102370.

Spenkuch, J. L. (2018). Expressive vs. strategic voters: An empirical assessment. *Journal of Public Economics 165*, 73–81.

Tausanovitch, C. and C. Warshaw (2018). Does the ideological proximity between candidates and voters affect voting in us house elections? *Political Behavior 40*, 223–245.

Thompson, D. M., J. A. Wu, J. Yoder, and A. B. Hall (2020). Universal vote-by-mail has no impact on partisan turnout or vote share. *Proceedings of the National Academy of Sciences 117*(25), 14052–14056.

Tomkins, S., K. Yao, J. Gaebler, T. Konitzer, D. Rothschild, M. Meredith, and S. Goel (2023). Blocks as geographic discontinuities: The effect of polling-place assignment on voting. *Political Analysis 31*(2), 165–180.

Uhlaner, C. J. (1989). Rational turnout: The neglected role of groups. *American Journal of Political Science 33*(2), 390–422.

Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political Science Review 77*, 142–157.

Yoder, J., C. Handan-Nader, A. Myers, T. Nowacki, D. M. Thompson, J. A. Wu, C. Yorgason, and A. B. Hall (2021). How did absentee voting affect the 2020 us election? *Science advances 7*(52), eabk1755.

Appendix A: Proofs

Proof of Lemma 1. First, note that $\gamma_L(y, S_L, S_R)$ is quadratic, single-peaked in y for $\theta \in (-\infty, \infty)$. Indeed,

$$\frac{\partial \gamma_L(y, S_L, S_R)}{\partial y} = 2\beta(x_L - y) - 2(1-\beta)\left(G(\Phi(E_L) - \Phi(S_R)) - G(\Phi(S_L) - \Phi(S_R))\right)(x_R - x_L).$$

Thus, (a) $\frac{\partial \gamma_L(y,S_L,S_R)}{\partial y}$ decreases linearly in y, (b) $\lim_{y\to-\infty} \frac{\partial \gamma_L(y,S_L,S_R)}{\partial y} > 0$ and (c) $\lim_{y\to\frac{x_L+x_R}{2}} \frac{\partial \gamma_L(y,S_L,S_R)}{\partial y} < 0$.

Now let $S_L \neq \emptyset$ and consider $\underline{y}, \bar{y} \in S_L$ with $\underline{y} < \bar{y}$. Consider any y between $\underline{y}$ and $\bar{y}$. For voting to be a best response for citizens $\underline{y}$ and $\bar{y}$, it must be $\gamma_L(\underline{y}, S_L, S_R) \geq 0$ and $\gamma_L(\bar{y}, S_L, S_R) \geq 0$. By the single-peakedness of $\gamma_L(y, S_L, S_R)$, we have $\gamma_L(y, S_L, S_R) > \min\{\gamma_L(\underline{y}, S_L, S_R), \gamma_L(\bar{y}, S_L, S_R)\} \geq 0$. Then, it must be that citizen y prefers to vote, and hence $y \in S_L$. Thus, S_L is an interval. The argument for S_R is analogous.

We next show that S_i does not contain points that are arbitrarily close to $\bar{x} = (x_L + x_R)/2$. Suppose by way of contradiction that there exists a sequence y_n, $n \in \mathbb{N}$ with $\lim_{n\to\infty} y_n = \bar{x}$ and $y_n \in S_L$ for all n. As $y_n \in S_L$, it must be that $\gamma_i(y_n, S_L, S_R) \geq 0$ for all n. Thus, continuity of γ_i implies that $0 \leq \lim_{n\to\infty} \gamma_i(y_n, S_L, S_R) = \gamma_i(\bar{x}, S_L, S_R) = \beta(A - (x_L - \bar{x})^2) - c$. Thus, we have must $A - (x_L - \bar{x})^2 \geq c$, which implies $\bar{x} - x_L \leq \sqrt{A - c}$. Substituting $\bar{x} = (x_L + x_R)/2$ and simplifying yields $2\sqrt{A - c} \geq x_R - x_L$ which contradicts the assumption that the midpoint between policies x_L and x_R is not contained in E_L. The argument for S_R is analogous.

To show closedness, let y_n be a sequence in S_L with $\lim_{n\to\infty} y_n = y$. As $y_n \in S_L$, we have $\gamma_L(y_n, S_L, S_R) \geq 0$. Hence, by continuity, $\gamma_L(y, S_L, S_R) \geq 0$, and thus $y \in S_L$. As we have shown that $y < \bar{x}$ and hence $y \notin S_R$, there is no other impediment for y to join coalition S_L. Closedness of S_R is established identically.

Finally, let y be a boundary point of S_i. Then, for voting to be a best response for y, we must have $\gamma_i(y, S_L, S_R) \geq 0$. Suppose by way of contradiction that $\gamma_i(y, S_L, S_R) > 0$. Then continuity of γ_i in y implies that there exists a $y' < \bar{x}$, $y' \notin S_i$, with $\gamma_i(y', S_L, S_R) > 0$. Then, the best response of citizen y' to coalitions (S_L, S_R) would be to vote, a contradiction. ■

Proof of Lemma 2. $\gamma_i(y, [y, \bar{y}_i], S_{-i})$ is continuously differentiable in y. Suppose by way of contradiction that $\frac{\partial}{\partial y}\gamma_i(y, [y, \bar{y}_i], S_{-i})_{y=\underline{y}_i} < 0$. For example, consider coalition S_L. Then there exists a $\delta > 0$ such that extending S_L to $S_L \cup [\underline{y}_L - \delta, \underline{y}_L)$ makes all new members strictly better off, contradicting the no minimal expansion condition. The arguments for expanding S_R to the left or S_L and S_R to the right are analogous. ■

Proof of Proposition 1. Substituting $S_L = -S_R$ into (1) and using symmetry of Φ yields

$$\tilde{\gamma}_L(y, S_L) := \gamma_L(y, S_L, S_R) = \beta(A - (x_L - y)^2) - c + 4(1-\beta)\left(G(\Phi(S_L \setminus E_L)) - \frac{1}{2}\right)x_L y.$$

The last summand is non-negative, and strictly positive if $\Phi(S_L) > \Phi(E_L)$. Thus, increasing S_L increases $\tilde{\gamma}_L(y, S_L)$.

Consider the set of all intervals $\mathcal{I}$ in $\mathbb{R}_-$ that contain E_L. Endow $\mathcal{I}$ with the natural order and define a mapping $T : \mathcal{I} \mapsto \mathcal{I}$ as follows. Let $[\underline{y}, \bar{y}] \in \mathcal{I}$ ($\underline{y}$ can be $-\infty$), and let $\underline{y}' < \underline{y}$ be a root of $\tilde{\gamma}_L(\underline{y}', [\underline{y}', \bar{y}]) = 0$ with $\tilde{\gamma}_L(y, [y, \bar{y}]) < 0$ for all $y \in (\underline{y}' - \delta, \underline{y}')$ for some $\delta > 0$. If $\tilde{\gamma}_L(\underline{y}', [\underline{y}', \bar{y}]) \geq 0$ for all $\underline{y}'$ then we define $\underline{y}' = -\infty$. Similarly, let $\bar{y}' > \bar{y}$ be a root of $\tilde{\gamma}_L(\bar{y}', [\underline{y}, \bar{y}']) = 0$ in $\mathbb{R}_-$ with $\tilde{\gamma}_L(y, [\underline{y}, y]) < 0$ for all $y \in (\bar{y}', 0)$. Then define $T([\underline{y}, \bar{y}]) = [\underline{y}', \bar{y}']$.

Next, note that T is monotone, i.e., if $S \subset S'$ then $T(S) \subset T(S')$. This follows because $\tilde{\gamma}_L$ increases as S is increased. Because $\mathcal{I}$ is a complete lattice, Tarski's fixed point theorem implies that there exists a maximal fixed point S^*. Because $T(S^*) = S^*$, voting is a best response for every citizen $y \in S^*$ (i.e., $\gamma_L(y, S_L, S_R) \geq 0$ for $S_L = S^*$ and $S_R = -S^*$.). Because S^* is maximal, abstaining is a best response for every potential candidate L supporter with $y \notin S^*$ (i.e., $\gamma_L(y, S_L, S_R) < 0$), and the equilibrium satisfies the no minimal expansion condition. The same result holds for $S_R = -S^*$. ■

Proof of Proposition 2. Consider (4). It is immediate that this equation is independent of S_R. It remains to prove that (a) a solution S_L exists or that $S_L = \emptyset$, (b) all voters in and out of S_L are best responding, and (c) the no minimal expansion condition is satisfied. This argument follows as in the proof of Proposition 1 by applying Tarski's fixed point theorem. ■

Proof of Proposition 3. If voters have purely selfish preferences, $\beta = 1$, then, regardless of the form of G and Φ, the net benefit of a citizen $y < (x_L + x_R)/2$ from having coalition S_L form is:

$$\gamma_L(y, S_L, S_R) = A - (x_L - y)^2 - c,$$

and the net benefit of a citizen $y > (x_L + x_R)/2$ from having coalition S_R form is:

$$\gamma_R(y, S_L, S_R) = A - (x_R - y)^2 - c.$$

Solving $A - (x_L - y)^2 - c = 0$, yields $\underline{y}_i = x_i - \sqrt{A - c}$ and $\bar{y}_i = x_i + \sqrt{A - c}$ for $i = L, R$. Our assumption that $x_R - x_L > 2\sqrt{A - c}$ ensures that $\bar{y}_L < (x_L + x_R)/2 < \underline{y}_R$. It is immediate that no other citizen wants to join the coalitions. Thus, $S_i = E_i = \left[x_i - \sqrt{A - c}, x_i + \sqrt{A - c}\right]$. ■

Proof of Proposition 4. Suppose that $S_L \neq \emptyset$. Then, voting being a best response for all $y \in S_L$ requires

$$\alpha\Phi(S_L)(x_R - x_L)(x_L + x_R - 2y) \geq c. \tag{11}$$

The left-hand side is decreasing in y. Thus, equality can only hold at a single y, and the inequality is strict for all $y' < y$. Hence, S_L takes the form $(-\infty, \bar{y}_L^*]$. Further, since abstaining must be a best response for $y > \bar{y}_L^*$, (11) must hold with equality at $y = \bar{y}_L^*$.

Consider equation (7). Note that $\gamma_L(\bar{y}_L, (\infty, \bar{y}_L], S_R) = -c/\alpha$ for $\bar{y}_L = \frac{x_L+x_R}{2}$. Further, $\gamma_L'(\bar{y}_L; S_R) = \phi(\bar{y}_L)(x_R - x_L)(x_L + x_R - 2\bar{y}_L) - 2(x_R - x_L)\Phi(\bar{y}_L)$. Thus, $\gamma_L'(\bar{y}_L; S_R) < 0$ if and only if

$$\frac{x_L + x_R}{2} < \frac{1}{\phi(\bar{y}_L)/\Phi(\bar{y}_L)} + \bar{y}_L. \tag{12}$$

(12) is clearly satisfied for $\bar{y}_L = \frac{x_L+x_R}{2}$. The right-hand side of (12) is increasing in $\bar{y}_L$ since $\phi(\bar{y}_L)/\Phi(\bar{y}_L)$ is decreasing by the log-concavity of Φ. Thus, (12) is violated, and hence $\gamma_L'(\bar{y}_L; S_R) > 0$, if $\bar{y}_L$ becomes sufficiently negative. Further, log concavity of Φ implies that γ_L' changes its sign once from positive to negative as $\bar{y}_L$ is increased towards $\frac{x_L+x_R}{2}$.

Let $\bar{c}_L = \max_{\bar{y}_L \in (-\infty, \frac{x_L+x_R}{2}]} \gamma_L(\bar{y}_L, S_L, S_R)$. For $c < \bar{c}_L$, there are two solutions, $\bar{y}_{1L} < \bar{y}_{2L}$, to the condition $\gamma_L(\bar{y}_L, (\infty, \bar{y}_L], S_R) = 0$ on $\bar{y}_L \in (-\infty, \frac{x_L+x_R}{2})$. Solution $\bar{y}_{1L}$ is not admissible, because $\gamma_L'(\bar{y}_{1L}; S_R) > 0$, violating Lemma 2. The unique equilibrium coalition is $S_L = (-\infty, \bar{y}_{2L}]$, because $\gamma_L'(\bar{y}_{2L}; S_R) < 0$. When $c = \bar{c}_L$, there is a unique $\bar{y}_L^* \in (-\infty, \frac{x_L+x_R}{2})$ such that $\gamma_L(\bar{y}_L^*, (\infty, \bar{y}_L^*], S_R) = 0$, and $\gamma_L'(\bar{y}_L^*; S_R) = 0$. The unique equilibrium coalition S_L is $(-\infty, \bar{y}_L^*]$. For $c > \bar{c}_L$, there does not exist any $\bar{y}_L^* \in (-\infty, \frac{x_L+x_R}{2})$ such that $\gamma_L(\bar{y}_L^*, (\infty, \bar{y}_L^*], S_R) = 0$, and hence the equilibrium coalition is empty: $S_L = \emptyset$.

The argument for S_R is analogous. ■

Proof of Proposition 5. Consider $S_L = [\underline{y}_L, \bar{y}_L]$ with $\underline{y}_L < \bar{y}_L$. Rewrite (4) as:

$$\beta(A - (x_L - y)^2) - c + (1-\beta)\left(G(\Phi(S_L) - \Phi(S_R)) - G(\Phi(E_L) - \Phi(S_R))\right) 2(x_R - x_L)\left(\frac{x_L + x_R}{2} - y\right) = 0$$

From $\underline{y}_L < \bar{y}_L < \frac{x_L+x_R}{2}$, we have $\frac{x_L+x_R}{2} - \underline{y}_L > \frac{x_L+x_R}{2} - \bar{y}_L$. This implies that the final term must be higher for $y = \underline{y}_L$ than for $y = \bar{y}_L$. Hence, the first term must be lower for $y = \underline{y}_L$ than for $y = \bar{y}_L$, that is to say $\left(x_L - \underline{y}_L\right)^2 > (x_L - \bar{y}_L)^2$, or $|\underline{y}_L - x_L| > |\bar{y}_L - x_L|$. Analogous arguments for S_R yield $|x_R - \bar{y}_R| > |x_R - \underline{y}_R|$. ■

Proof of Proposition 6. Recall that if $S_L = [\underline{y}_L, \bar{y}_L]$, then $\underline{y}_L$ and $\bar{y}_L$ satisfy (4). Proposition 4 shows that $\gamma(y, (-\infty, y], S_R)$ is strictly single peaked when $\beta = 0$ if $c < \bar{c}_L$. Because γ is continuously differentiable in β it follows that $\gamma(y, [\underline{y}, y], S_R)$ is also single peaked in y for β close to 0 and $\underline{y}$ sufficiently negative.

Next, note that $E_L = \emptyset$ near $\beta = 0$. Thus, $\Phi(E_L) = 0$, which implies that

$$(1-\beta)\alpha\left(\Phi(S_L) - \Phi(E_L)\right)\left((x_L - y)^2 - (x_R - y)^2\right) \tag{13}$$

is decreasing in β near $\beta = 0$. It is also immediate that $-\beta(x_L - y)^2$ is decreasing in β. Hence $\gamma(y, [\underline{y}, y], S_R)$ is decreasing in β if and only if A is not too large. Further, if $\gamma(y, [\underline{y}, y], S_R)$ is decreased, then, as Figure 3 illustrates, $\bar{y}_L$ must decrease to retrieve equality in equation (4).

Next, consider the equilibrium equation for $\underline{y}_L$. The same argument as above implies that (13) is decreasing in β. Further, because $\underline{y}_L = -\infty$ at $\beta = 0$, the term $A - (x_L - \underline{y}_L)^2 < 0$ for small β. Hence, increasing β (near $\beta = 0$) lowers the γ function and hence increases $\underline{y}_L$. Thus, $\Phi(S_L)$ is decreasing in β close to zero for A sufficiently small.

As we increase A, the derivative with respect to β increases. Thus, there exists a critical value $\bar{A}$ such that $\Phi(S_L)$ and the turnout rate are decreasing for $A < \bar{A}$ and increasing for $A > \bar{A}$. The argument for S_R is analogous.

For $\beta \to 1$, it must be that $S_L \to E_L$. We can approximate the comparative statics by assuming that $S_L = E_L = \left[x_i - \sqrt{A - \frac{c}{\beta}}, x_i + \sqrt{A - \frac{c}{\beta}}\right]$. Clearly, $\Phi(E_L)$ is increasing in β as the coalition expands on both sides. Thus, for β close to 1, turnout is increasing in β. We also have that the turnout rate $r_L = \frac{\Phi(E_L)}{\Phi(\frac{x_L + x_R}{2})}$ increases in β since $\frac{x_L + x_R}{2}$ is unaffected by β. ■

Proof of Proposition 7. First consider $\beta = 0$. Differentiating the equilibrium condition $\gamma_L(\bar{y}_L(\mu), (\infty, \bar{y}_L(\mu)], S_R) = 0$ with respect to μ at $\mu = 0$ with symmetrically-situated candidates, $x_R = -x_L$, yields

$$\Phi(\bar{y}_L(\mu)) \frac{\partial \bar{y}_L(\mu)}{\partial \mu}\Big|_{\mu=0} + \phi(\bar{y}_L(\mu)) \left(\frac{\partial \bar{y}_L(\mu)}{\partial \mu}\Big|_{\mu=0} - 1\right) \bar{y}_L(\mu) = 0,$$

which we solve for

$$\frac{\partial \bar{y}_L(\mu)}{\partial \mu}\Big|_{\mu=0} = 1 - \frac{\Phi(\bar{y}_L)}{\Phi(\bar{y}_L) + \bar{y}_L \phi(\bar{y}_L)}. \tag{14}$$

Result 4 implies that $\bar{y}_L$ is strictly decreasing in μ if $c < \bar{c}_L = \bar{c}_R$, where $\bar{c}_i$ is defined in (9). Thus,

$$\Phi(\bar{y}_L) + \bar{y}_L \phi(\bar{y}_L) > 0. \tag{15}$$

Differentiating $r_L = \Phi(\bar{y}_L - \mu) / \Phi(-\mu)$ with respect to μ at $\mu = 0$ and denoting $\partial \bar{y}_L(\mu) / \partial \mu_{|_{\mu=0}}$ by $\bar{y}'_L$ yields

$$\frac{\partial r_L}{\partial \mu}\Big|_{\mu=0} = \frac{\phi(\bar{y}_L)\left(\bar{y}'_L - 1\right)\Phi(0) + \phi(0)\Phi(\bar{y}_L)}{\Phi(0)^2}.$$

Plugging the expression for $\bar{y}'_L$ from (14), we obtain

$$\frac{\partial r_L}{\partial \mu}|_{\mu=0} = \frac{\phi(\bar{y}_L)\left(-\frac{\Phi(\bar{y}_L)}{\Phi(\bar{y}_L) + \bar{y}_L \phi(\bar{y}_L)}\right)\Phi(0) + \phi(0)\Phi(\bar{y}_L)}{\Phi(0)^2}$$

which has the same sign as

$$-\phi(\bar{y}_L)\left(\frac{\Phi(y_L)}{\Phi(\bar{y}_L)+\bar{y}_L\phi(\bar{y}_L)}\right)\Phi(0)+\phi(0)\Phi(\bar{y}_L)=\Phi(\bar{y}_L)\left(\phi(0)-\frac{1}{2}\left(\frac{\phi(\bar{y}_L)}{\Phi(\bar{y}_L)+\bar{y}_L\phi(\bar{y}_L)}\right)\right). \tag{16}$$

Given $\Phi(\bar{y}_L)+\bar{y}_L\phi(\bar{y}_L)>0$, the sign of (16) does not change when multiplied by $\Phi(\bar{y}_L)+y_L\phi(\bar{y}_L)$, which yields

$$\Phi(\bar{y}_L)\left(\phi(0)(\Phi(\bar{y}_L)+\bar{y}_L\phi(\bar{y}_L))-\frac{1}{2}\phi(\bar{y}_L)\right). \tag{17}$$

The sign of (17) is the same as

$$\phi(0)\Phi(\bar{y}_L)+\phi(\bar{y}_L)\left(\bar{y}_L\phi(0)-\frac{1}{2}\right)=\Phi(\bar{y}_L)\left(\phi(0)+\frac{\phi(\bar{y}_L)}{\Phi(\bar{y}_L)}\left(\bar{y}_L\phi(0)-\frac{1}{2}\right)\right).$$

Because $\bar{y}_L<0$, the log concavity of Φ yields $\phi(\bar{y}_L)/\Phi(\bar{y}_L)\geq\phi(0)/\Phi(0)$. Thus, given $\bar{y}_L\phi(0)-\frac{1}{2}<0$ (as $\bar{y}_L<0$), we have

$$\begin{aligned}\Phi(\bar{y}_L)\left(\phi(0)+\frac{\phi(\bar{y}_L)}{\Phi(\bar{y}_L)}\left(\bar{y}_L\phi(0)-\frac{1}{2}\right)\right)&\leq\Phi(\bar{y}_L)\left(\phi(0)+\frac{\phi(0)}{\Phi(0)}\left(\bar{y}_L\phi(0)-\frac{1}{2}\right)\right)\\&=\Phi(\bar{y}_L)\left(\phi(0)+2\phi(0)\left(\bar{y}_L\phi(0)-\frac{1}{2}\right)\right)\\&=\Phi(\bar{y}_L)\,2\bar{y}_L\phi(0)^2<0.\end{aligned}$$

Thus, $\frac{\partial r_L}{\partial\mu}|_{\mu=0}<0$, i.e., the participation rate, r_L, of L supporters decreases as μ is increased. Conversely, decreasing μ increases the turnout rate for candidate L. The result for candidate R follows by symmetry.

Finally, suppose that $\beta=1$. It is immediate that $\Phi_{\mu,\sigma}(E_L)$ decreases as μ is increased. Let $E_L=[\underline{y},\bar{y}]$. The turnout rate is

$$r_L=\frac{\Phi\left(\frac{\bar{y}-\mu}{\sigma}\right)-\Phi\left(\frac{\underline{y}-\mu}{\sigma}\right)}{\Phi\left(\frac{-\mu}{\sigma}\right)}.$$

Then

$$\left.\frac{\partial r_L}{\partial\mu}\right|_{\mu=0}=\frac{4\phi(0)\left(\Phi\left(\frac{\bar{y}}{\sigma}\right)-\Phi\left(\frac{\underline{y}}{\sigma}\right)\right)-2\left(\phi\left(\frac{\bar{y}}{\sigma}\right)-\phi\left(\frac{\underline{y}}{\sigma}\right)\right)}{\sigma}=\frac{4\phi(0)-2\frac{\phi\left(\frac{\bar{y}}{\sigma}\right)-\phi\left(\frac{\underline{y}}{\sigma}\right)}{\Phi\left(\frac{\bar{y}}{\sigma}\right)-\Phi\left(\frac{\underline{y}}{\sigma}\right)}}{\sigma\left(\Phi\left(\frac{\bar{y}}{\sigma}\right)-\Phi\left(\frac{\underline{y}}{\sigma}\right)\right)}. \tag{18}$$

L'Hôpital's rule implies that

$$\lim_{\sigma\to\infty}\frac{\phi\left(\frac{\bar{y}}{\sigma}\right)-\phi\left(\frac{\underline{y}}{\sigma}\right)}{\Phi\left(\frac{\bar{y}}{\sigma}\right)-\Phi\left(\frac{\underline{y}}{\sigma}\right)}=\lim_{\sigma\to\infty}\frac{-\frac{\bar{y}}{\sigma^2}\phi'\left(\frac{\bar{y}}{\sigma}\right)+\frac{\underline{y}}{\sigma^2}\phi'\left(\frac{\underline{y}}{\sigma}\right)}{-\frac{\bar{y}}{\sigma^2}\phi\left(\frac{\bar{y}}{\sigma}\right)+\frac{\underline{y}}{\sigma^2}\phi\left(\frac{\underline{y}}{\sigma}\right)}=\lim_{\sigma\to\infty}\frac{-\bar{y}\phi'\left(\frac{\bar{y}}{\sigma}\right)+\underline{y}\phi'\left(\frac{\underline{y}}{\sigma}\right)}{-\bar{y}\phi\left(\frac{\bar{y}}{\sigma}\right)+\underline{y}\phi\left(\frac{\underline{y}}{\sigma}\right)}=0, \tag{19}$$

because Φ is single peaked and symmetric so $\phi'(0)=0$, and the denominator converges to $\phi(0)(\underline{y}-\bar{y})\neq0$. Thus, the numerator of (18) becomes strictly positive, and $\frac{\partial r_L}{\partial\mu}|_{\mu=0}>0$. ■

Proof of Proposition 8. For $c = 0$ and $\beta = 1$ Proposition 3 implies that $E_i = \left[x_i - \sqrt{A}, x_i + \sqrt{A}\right]$ if $x_i \geq \sqrt{A}$. Candidate L's winning probability is therefore $G(\Phi(E_L) - \Phi(E_R))$. The first-order condition for $\beta = 1$ and $c = 0$ for candidate R and evaluating at $x_L = -x_R$ yields

$$-4g(0)\left(\phi\left(x_R - \sqrt{A}\right) - \phi\left(x_R + \sqrt{A}\right)\right)x_R\theta_R + \theta_R - x_R = 0,$$

which implies

$$x_R = \frac{\theta_R}{1 + 4g(0)\left(\phi\left(x_R - \sqrt{A}\right) - \phi\left(x_R + \sqrt{A}\right)\right)\theta_R}. \tag{20}$$

If $x_R < \sqrt{A}$ then $E_R = [0, x_R + \sqrt{A}]$ and the first-order condition becomes

$$4g(0)\phi\left(x_R + \sqrt{A}\right)x_R\theta_R + \theta_R - x_R = 0,$$

which cannot be satisfied because the LHS is strictly possible. This means that it is optimal for candidate R to move past $\sqrt{A}$, which implies that the equilibrium x_R is given by (20)

In contrast, if $\beta = 0$ then (7) and (8) yield $S_L = (-\infty, \bar{x}]$ and $S_R = [\bar{x}, \infty)$, where $\bar{x} = (x_L + x_R)/2$. Candidate L's winning probability is therefore $G(2\Phi(\bar{x}) - 1)$. The first-order condition for candidate R taken at $x_R = -x_L$ is therefore

$$-4g(0)\phi(0)x_R\theta_R + \theta_R - x_R = 0,$$

which implies

$$x_R = \frac{\theta_R}{1 + 4g(0)\phi(0)\theta_R}. \tag{21}$$

Because Φ is single peaked, $\phi(0) > \phi\left(x_R - \sqrt{A}\right) - \phi\left(x_R + \sqrt{A}\right)$ and hence (21) is strictly smaller than (20). Given that $x_L = -x_R$ in both cases, this implies that elite polarization is lower with purely instrumental preferences when $c = 0$. The monotonicity result in Proposition 9 concludes the proof that elite polarization is strictly smaller for social preferences when $c < \bar{c}$. ■

Proof of Proposition 9. *Purely cooperative preferences: $\beta = 0$.*

The coalitions are given by $S_L = (-\infty, \bar{y}_L]$ and $S_R = [\underline{y}_R, +\infty)$, where $\bar{y}_L$ and $\underline{y}_R$ satisfy (7) and (8). Hence, $\underline{y}_R$ solves:

$$\alpha(1 - \Phi(\underline{y}_R))\left(x_R - x_L\right)\left(2\underline{y}_R - x_L - x_R\right) - c = 0,$$

or, using $x_L + x_R = 0$,

$$4\alpha x_R\underline{y}_R(1 - \Phi(\underline{y}_R)) - c = 0. \tag{22}$$

Here,

$$\left.\frac{\partial\bar{y}_L}{\partial x_R}\right|_{-x_L = x_R} = -\left.\frac{\partial\gamma_L(\bar{y}_L, S_L, S_R)/\partial x_R}{\partial\gamma_L(\bar{y}_L, S_L, S_R)/\partial\bar{y}_L}\right|_{-x_L = x_R} = \frac{1}{2x_R}\frac{(1 - \Phi(\underline{y}_R))(x_R + \underline{y}_R)}{1 - \Phi(\underline{y}_R) - \underline{y}_R\phi\left(\underline{y}_R\right)}$$

$$\left.\frac{\partial \underline{y}_R}{\partial x_R}\right|_{-x_L=x_R} = -\left.\frac{\partial \gamma_R(\underline{y}_R, S_L, S_R)/\partial x_R}{\partial \gamma_R(\underline{y}_R, S_L, S_R)/\partial \underline{y}_R}\right|_{-x_L=x_R} = \frac{1}{2x_R}\frac{(1-\Phi(\underline{y}_R))(x_R-\underline{y}_R)}{1-\Phi(\underline{y}_R)-\underline{y}_R\phi\left(\underline{y}_R\right)}$$

By Lemma 2, the admissible solutions $\bar{y}_L$ and $\underline{y}_R$ must be such that $\gamma'_L(y; S_R) \leq 0$ and $\gamma'_R(y; S_L) \geq 0$. Hence, we have $x_R\left[\underline{y}_R\phi(\underline{y}_R) - (1-\Phi(\underline{y}_R))\right] \leq 0$, i.e., $\underline{y}_R\phi(\underline{y}_R) \leq 1-\Phi(\underline{y}_R)$. Further,

$$\left.\frac{\partial \bar{y}_L}{\partial x_R}\right|_{-x_L=x_R} \geq \left.\frac{\partial \underline{y}_R}{\partial x_R}\right|_{-x_L=x_R} \geq 0,$$

and

$$\left.\frac{\partial \bar{y}_L}{\partial x_R}\right|_{-x_L=x_R} + \left.\frac{\partial \underline{y}_R}{\partial x_R}\right|_{-x_L=x_R} = \frac{1-\Phi(\underline{y}_R)}{1-\Phi(\underline{y}_R)-\underline{y}_R\phi\left(\underline{y}_R\right)}$$

Candidate R solves

$$\max_{x_R}\left\{EU_R(x_R, x_L) = -\left[\frac{1}{2}+\alpha\left[\Phi(\bar{y}_L)-(1-\Phi(\underline{y}_R))\right]\right](x_L-\theta_R)^2 - \left[\frac{1}{2}-\alpha\left[\Phi(\bar{y}_L)-(1-\Phi(\underline{y}_R))\right]\right](x_R-\theta_R)^2\right\}$$

$$\begin{aligned}\frac{\partial}{\partial x_R}EU_R(x_R, x_L) &= -(x_R-\theta_R)+2\alpha\left[\Phi(\bar{y}_L)-(1-\Phi(\underline{y}_R))\right](x_R-\theta_R)\\ &\quad -\alpha\left[\phi(\bar{y}_L)\frac{\partial \bar{y}_L}{\partial x_R}+\phi(\underline{y}_R)\frac{\partial \underline{y}_R}{\partial x_R}\right]\left[(x_L-\theta_R)^2-(x_R-\theta_R)^2\right]\end{aligned}$$

The FOC is:

$$\begin{aligned}\theta_R - x_R &= 4x_R\theta_R\alpha\phi(\underline{y}_R)\left(\frac{\partial \bar{y}_L}{\partial x_R}+\frac{\partial \underline{y}_R}{\partial x_R}\right) \qquad (23)\\ &= 4\alpha x_R\theta_R\phi(\underline{y}_R)\frac{1-\Phi(\underline{y}_R)}{1-\Phi(\underline{y}_R)-\underline{y}_R\phi\left(\underline{y}_R\right)}\end{aligned}$$

or

$$(\theta_R - x_R)\left(1-\frac{\underline{y}_R\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)}\right) = 4\alpha x_R\theta_R\phi(\underline{y}_R) \qquad (24)$$

Given $\underline{y}_R\phi(\underline{y}_R) \leq 1-\Phi(\underline{y}_R)$, $1-\frac{\underline{y}_R\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)} > 0$.

Again, the left-hand side is the marginal benefit for diverging from zero close to θ_R, whereas the right-hand side is the marginal cost, in terms of probability of losing the election $\alpha\phi(\underline{y}_R)\left(\frac{\partial \bar{y}_L}{\partial x_R}+\frac{\partial \underline{y}_R}{\partial x_R}\right)$ times the cost if losing, $(x_L-\theta_R)^2-(x_R-\theta_R)^2 = 4x_R\theta_R$.

We want the implicit function theorem to imply that the symmetric equilibrium platform x_R increases in c and θ_R for all c sufficiently small. Solve for x_R in equation (22),

$$x_R = \frac{c}{4\alpha\underline{y}_R(1-\Phi(\underline{y}_R))},$$

and substitute into equation (24) to obtain:

$$\lambda(\underline{y}_R, c) \equiv \left(\theta_R - \frac{c}{4\alpha \underline{y}_R(1-\Phi(\underline{y}_R))}\right)\left(1 - \underline{y}_R \frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)}\right) - c\frac{\theta_R \phi(\underline{y}_R)}{\underline{y}_R(1-\Phi(\underline{y}_R))} = 0. \tag{25}$$

For small c,

$$\frac{d\underline{y}_R}{dc} = -\frac{\partial\lambda(\underline{y}_R, c)/\partial c}{\partial\lambda(\underline{y}_R, c)/\partial \underline{y}_R} = -\frac{-\frac{1-\underline{y}_R\frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)}}{4\alpha \underline{y}_R(1-\Phi(\underline{y}_R))} - \frac{\theta_R\phi(\underline{y}_R)}{\underline{y}_R(1-\Phi(\underline{y}_R))}}{\theta_R \frac{d}{d\underline{y}_R}\left(1 - \underline{y}_R\frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)}\right)} = -\frac{-\frac{1-\underline{y}_R\frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)}}{4\alpha \underline{y}_R(1-\Phi(\underline{y}_R))} - \frac{\theta_R\phi(\underline{y}_R)}{\underline{y}_R(1-\Phi(\underline{y}_R))}}{-\frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)} - \underline{y}_R\frac{d}{d\underline{y}_R}\frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)}} < 0,$$

from our increasing hazard rate condition, $\frac{d}{d\underline{y}_R}\frac{\phi\left(\underline{y}_R\right)}{1-\Phi(\underline{y}_R)} > 0$.

Totally differentiating (22) yields

$$4\alpha \underline{y}_R(1-\Phi(\underline{y}_R))\frac{dx_R}{dc} + 4\alpha x_R\left(1-\Phi(\underline{y}_R)\right)\left(1 - y_R\frac{\phi\left(y_R\right)}{1-\Phi(\underline{y}_R)}\right)\frac{dy_R}{dc} = 1,$$

and because $4\alpha \underline{y}_R(1-\Phi(\underline{y}_R)) > 0$, $4\alpha x_R\left(1-\Phi(\underline{y}_R)\right)\left(1 - y_R\frac{\phi(y_R)}{1-\Phi(\underline{y}_R)}\right) > 0$ and $\frac{dy_R}{dc} < 0$, it follows that $\frac{dx_R}{dc} > 0$. Thus, the platform x_R increases in c.

Purely selfish preferences: $\beta = 1$. With $x_L = -x_R$, $\frac{x_L+x_R}{2} = 0$. For $\beta = 1$, the sets of voters who participate takes the form

$$E_L = [x_L - z, \min\{x_L + z, 0\}], \text{ and } E_R = [\max\{x_R - z, 0\}, x_R + z],$$

where $z \equiv \sqrt{A-c}$. Suppose first that $z \geq x_R$. Then candidate L wins if $\Phi(0)-\Phi(x_L-z) \geq \Phi(x_R+z)-\Phi(0)+\rho$, where ρ is the realized net partisan advantage for candidate R. Thus, candidate L's winning probability is

$$P(x_L, x_R) = G(2\Phi(0) - \Phi(x_L - z) - \Phi(x_R + z)).$$

Candidate R solves

$$\max_{x_R} -P(x_L, x_R)(x_L - \theta_R)^2 - (1 - P(x_L, x_R))(x_R - \theta_R)^2 \tag{26}$$

Let $EU_R(x_R, x_L)$ be the value of the objective.

Taking the first order condition at $x_L = -x_R$ yields

$$\begin{aligned}
\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L=x_R} &= -\frac{\partial}{\partial x_R}P(x_L, x_R)[(x_L - \theta_R)^2 - (x_R - \theta_R)^2] - 2(1 - P(x_L, x_R))(x_R - \theta_R)\Big|_{-x_L=x_R} \\
&= -g(2\Phi(0) - \Phi(x_L - z) - \Phi(x_R + z))|_{-x_L=x_R}\left[2\phi(0) - \phi(x_R + z)\right]4\theta_R x_R - (x_R - \theta_R) \\
&= -4g(0)\left[2\phi(0) - \phi(x_R + z)\right]\theta_R x_R - (x_R - \theta_R) = 0,
\end{aligned} \tag{27}$$

using $\Phi(0) - \Phi(x_L - z) = \Phi(x_R + z) - \Phi(0)$ implied by the symmetry of Φ at 0.

In words, the marginal benefit $\theta_R - x_R$ for raising x_R closer to θ_R must be met by the marginal cost, which is given by the marginal increase $g(0)\left[\phi(0) - \phi(x_R + z)\right]$ in the risk of losing the election times the consequent utility loss $(x_L - \theta_R)^2 - (x_R - \theta_R)^2 = 4\theta_R x_R$.

Differentiating the left-hand side of (27) with respect to x_R yields:

$$\frac{\partial}{\partial x_R}\left(\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L = x_R}\right) = -4g(0)\left[2\phi(0) - \phi(x_R + z)\right] x_R - 4g(0)\left[-\phi'(x_R + z)\right]\theta_R x_R - 1 < 0,$$

where the inequality follows from the single-peakedness of Φ, which implies that $\phi'(x_R + z) < 0$.

Differentiation of equation (27) with respect to θ_R and z and the implicit function theorem imply that x_R decreases in z, whereas x_R increases in θ_R if $g(0)$ and $\phi(0)$ are sufficiently small, because

$$\frac{\partial}{\partial z}\left(\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L = x_R}\right) = -4g(0)\left[-\phi'(x_R + z)\right]\theta_R x_R < 0,$$

$$\frac{\partial}{\partial \theta_R}\left(\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L = x_R}\right) = 1 - 4g(0)\left[2\phi(0) - \phi(x_R + z)\right] x_R.$$

Now consider the case where $x_L + z < 0[= 0] < x_R - c$. Candidate L wins if $\Phi(x_L + z) - \Phi(x_L - z) \geq \Phi(x_R + z) - \Phi(x_R - z) + \rho$, where ρ is the realized net partisan advantage for candidate R. Thus, candidate L's winning probability is

$$P(x_L, x_R) = G(\Phi(x_L + z) - \Phi(x_L - z) + \Phi(x_R - z) - \Phi(x_R + z)). \tag{28}$$

Accounting for this difference, candidate R's optimization problem is the same as above. The first-order condition at $x_L = -x_R$ is now:

$$\begin{aligned}
&\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L = x_R} = -\frac{\partial}{\partial x_R} P(x_L, x_R)[(x_L - \theta_R)^2 - (x_R - \theta_R)^2] - 2(1 - P(x_L, x_R))(x_R - \theta_R)\Big|_{-x_L = x_R} \\
= \;& -g(\Phi(x_L + z) - \Phi(x_L - z) + \Phi(x_R - z) - \Phi(x_R + z))|_{-x_L = x_R}\left[\phi(x_R - z) - \phi(x_R + z)\right] 4\theta_R x_R - (x_R - \theta_R) \\
= \;& -4g(0)\left[\phi(x_R - z) - \phi(x_R + z)\right]\theta_R x_R - (x_R - \theta_R) = 0,
\end{aligned} \tag{29}$$

because $-x_L = x_R$ and symmetry of Φ at zero imply $\Phi(x_R + z) = \Phi(-x_L + z) = 1 - \Phi(x_L - z)$ and $\Phi(x_R - z) = \Phi(-x_L - z) = 1 - \Phi(x_L + z)$.

Differentiating the left-hand side of (29) with respect to x_R now gives:

$$\frac{\partial}{\partial x_R}\left(\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L = x_R}\right) = -4g(0)\theta_R[\left[\phi'(x_R - z) - \phi'(x_R + z)\right] x_R + \phi(x_R - z) - \phi(x_R + z)] - 1 < 0,$$

as long as $\left[\phi'(x_R - z) - \phi'(x_R + z)\right] x_R + \phi(x_R - z) - \phi(x_R + z) > 0$. Noting that $\phi(x_R - z) > \phi(x_R + z)$ by single-peakedness, a sufficient condition is that $\phi'(x_R - z) > \phi'(x_R + z)$.

Differentiating (29) with respect to θ_R and z and the implicit function theorem imply again that x_R decreases in z, whereas x_R increases in θ_R if $g(0)$ and $\phi(0)$ are sufficiently small:

$$\frac{\partial}{\partial \theta_R}\left(\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L=x_R}\right) = -4g(0)\left[\phi(x_R - z) - \phi(x_R + z)\right] x_R + 1$$

$$\frac{\partial}{\partial z}\left(\frac{\partial}{\partial x_R} EU_R(x_R, x_L)\Big|_{-x_L=x_R}\right) = 4g(0)\left[\phi'(x_R - z) + \phi'(x_R + z)\right] \theta_R x_R < 0.$$

To verify that the second-order condition holds, note that the second derivative of candidate R's optimization problem (26) is

$$-\alpha \frac{\partial^2 P}{\partial x_R^2}\left((x_L - \theta_R)^2 - (x_R - \theta_R)^2\right) - 4\alpha \frac{\partial P}{\partial x_R}(\theta_R - x_R) - 1. \tag{30}$$

Note that

$$\frac{\partial P}{\partial x_R} = \alpha\left(\phi(x_R - z) - \phi(x_R + z)\right) > 0 \text{ and } \frac{\partial^2 P}{\partial x_R^2} = \alpha\left(\phi'(x_R - z) - \phi'(x_R + z)\right).$$

Hence the second term in (30) is strictly negative. Further, the first term is negative if the density is concave between θ_L and θ_R. Alternatively, the first term becomes irrelevant if the density is sufficiently flat. ■

Appendix B: Additional Comparative Statics

Using Propositions 3 and 4, we can investigate the impact of mass polarization. Define $\Phi_s = \Phi(x/s)$. If Φ has standard deviation σ then Φ_s has standard deviation $s\sigma$. Increasing s corresponds to increasing the dispersion of citizen ideologies and hence increases mass polarization.

Result 5 *Suppose $x_L = -x_R$ and Φ is symmetric and mass polarization is increased by replacing $\Phi(x)$ with $\Phi(x/s)$ where $s > 1$. For $\beta = 0$ this causes S_i to increase, as well as both $\Phi(S_i)$ and r_i to increase (for $c_i < \bar{c}_i$, where $\bar{c}_i$ is given in Proposition 4). For $\beta = 1$, this does not affect E_i, but the effect on $\Phi(E_i)$ and r_i depends on parameters.*

With symmetry $\xi_{R,s}(y) = 4x_R y(1 - \Phi(y/s))$. It is immediate that $\xi_{R,s}$ is increasing in s for $y > 0$, so we can again conclude that $\bar{y}_R$ decreases and hence the interval of types S_R increases. Further, the number of centrist citizens decreases, i.e., $1 - \Phi_s(y)$ increases as s is increased, and hence $1 - \Phi(S_R)$ increases. Because the number of R supporters does not change, the turnout rate rises. For $\beta = 1$ it is immediate that the sets E_L and E_R do not change. We show in Lemma 3 below that turnout and turnout rates increase are single-peaked functions of mass polarization s, increasing when s is small and decreasing when s is large.

Lemma 3 *Suppose that $x\phi'(x)/\phi(x)$ is decreasing in x, which holds for normal distributions. Let $\Phi_s(x) = \Phi(x/s)$. Then there exists $\bar{s} \geq 0$ such that turnout decreases for $s > \bar{s}$, and increases for $s < \bar{s}$ as mass polarization, s, is increased when $\beta = 1$. Further, $\bar{s} > 0$ if Φ is a normal distribution.*

Proof of Lemma 3. Let $S_R = [\underline{y}, \bar{y}]$. The measure of voters for candidate R is $\Phi(\bar{y}/s) - \Phi(\underline{y}/s)$. The derivative with respect to s is

$$\frac{\underline{y}}{s^2}\phi\left(\frac{\underline{y}}{s}\right) - \frac{\bar{y}}{s^2}\phi\left(\frac{\bar{y}}{s}\right) = \frac{\phi\left(\frac{\underline{y}}{s}\right)}{s^2}\left(\underline{y} - \bar{y}\frac{\phi\left(\frac{\bar{y}}{s}\right)}{\phi\left(\frac{\underline{y}}{s}\right)}\right). \tag{31}$$

Note that

$$\lim_{s\to\infty} \underline{y} - \bar{y}\frac{\phi\left(\frac{\bar{y}}{s}\right)}{\phi\left(\frac{\underline{y}}{s}\right)} = \underline{y} - \bar{y} < 0. \tag{32}$$

Further

$$\frac{\partial}{\partial s}\left(\underline{y} - \bar{y}\frac{\phi\left(\frac{\bar{y}}{s}\right)}{\phi\left(\frac{\underline{y}}{s}\right)}\right) = \frac{\bar{y}}{s^2}\frac{\bar{y}\phi'\left(\frac{\bar{y}}{s}\right)\phi\left(\frac{\underline{y}}{s}\right) - \underline{y}\phi'\left(\frac{\underline{y}}{s}\right)\phi\left(\frac{\bar{y}}{s}\right)}{\phi^2\left(\frac{\underline{y}}{s}\right)} = \frac{\bar{y}\phi\left(\frac{\bar{y}}{s}\right)\phi\left(\frac{\underline{y}}{s}\right)}{s\phi^2\left(\frac{\underline{y}}{s}\right)}\left(\frac{\bar{y}}{s}\frac{\phi'\left(\frac{\bar{y}}{s}\right)}{\phi\left(\frac{\underline{y}}{s}\right)} - \frac{\underline{y}}{s}\frac{\phi'\left(\frac{\underline{y}}{s}\right)}{\phi\left(\frac{\bar{y}}{s}\right)}\right). \tag{33}$$

Because $x\phi'(x)/\phi(x)$ is decreasing in x and $\bar{y} > \underline{y}$, it follows that (33) is negative. This, in turn, implies that the sign of (31) can change at most once, i.e., if the sign of (31) is negative for small s then it is always negative. Otherwise, it becomes zero for some $\bar{s}$ and negative for $s > \bar{s}$.

Finally, note that for normal distributions

$$\lim_{s\to 0} \underline{y} - \bar{y}\frac{\phi\left(\frac{\bar{y}}{s}\right)}{\phi\left(\frac{\underline{y}}{s}\right)} = \underline{y} > 0. \tag{34}$$

Thus, there exists a unique value $\bar{s} > 0$ for which (31) becomes zero. ■