ABSTRACT

Title of dissertation: ESSAYS ON ELECTORAL MODELS

Emekcan Yucel, Doctor of Philosophy, 2017

Dissertation directed by: Professor Allan Drazen
Department of Economics

This dissertation studies how people vote and how politicians maximize votes. In the first chapter, I propose non-instrumental benefits to sincere voting as the explanation for why people vote for candidates certain to lose in elections. Building on this idea, I provide a framework where the decision of whether to vote sincerely or strategically is an endogenous choice that responds to election-specific characteristics, rather than a characteristic of a voter. Using both pivotal voter and group rule-utilitarian frameworks, I show that third party vote shares are lower and the extent of strategic voting is higher when the election is expected to be close or when the stakes of the election are high. I also show that adding a heterogeneous non-instrumental sincere voting benefit implies partial strategic desertion of weak parties by their supporters and a lower participation rate for minor party supporters compared to major party supporters. Furthermore, I present theoretical predictions on the impact of electorate size on third party vote shares and on the correlation between third party voting and turnout. Using data from U.S. presidential elections between 1920 and 2012, I also present empirical evidence consistent with the theo-

retical predictions of this chapter. In the second chapter (joint with Professor Allan Drazen), we ask what the successful electoral strategies are and whether candidates should try to persuade "swing" voters or mobilize their "base". We present a model that can address these and related questions in a single unified framework. We relate electoral strategies to the characteristics of voting groups, with the answers to these questions sometimes being surprising. We show how a candidate may have different ways of winning for given characteristics of the electoral population, with possible "discontinuities" in electoral positions that win elections. We believe that the model we present helps clarify some key issues as well as presenting insights into some real-world experience.

ESSAYS ON ELECTORAL MODELS

by

Emekcan Yucel

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor Allan Drazen, Chair/Advisor
Professor Ethan Kaplan
Professor Erkut Ozbay
Professor Peter Murrell
Professor Piotr Swistak

Acknowledgments

I am grateful to my advisor Allan Drazen for his invaluable guidance and support throughout my Ph.D. studies. It was a great pleasure to be his student, research assistant and co-author, and his wisdom helped me much along the way.

I am indebted to Ethan Kaplan for his invaluable mentoring and guidance. I much benefited from his insightful feedback both in our personal meetings and in the Political Economy research group meetings that he organized.

I thank Erkut Ozbay for his valuable comments and support, John Shea for his detailed and useful feedback to the drafts of my work, Peter Murrell and Piotr Swistak for agreeing to serve on my committee, and the participants of our Political Economy research group for their useful comments and suggestions.

I am grateful to my mother Fatma and to my brother Kuzey for their endless love and support, and to Halime, whose company brought utmost joy to my life. Finally, a special thanks to my father Cemal, whose wisdom and intellectual curiosity inspired me to pursue a life in academia.

Table of Contents

List of Tables

List of Figures

Chapter 1: Third Party Voting: Vote One's Heart or One's Mind?

1.1 Introduction

When only a simple majority of votes are required to win an election, small parties that have no chance of winning the election potentially play an important role. These minor parties have the potential to affect the outcome of the election by altering the vote shares of likely winners, a phenomenon called the "spoiler" effect. Spoilers have historically proved important in terms of changing the outcome of simple plurality elections, as there are many elections in which third party candidates receive higher votes than the difference between contenders. Recent examples from U.S. presidential elections include Ross Perot and Ralph Nader. In 1992, Perot had a vote share that was much higher than the difference between Clinton and Bush. Many people believe that the presence of Perot denied victory to Bush since the majority of Perot voters would have voted for Bush had Perot not run. Likewise, it is widely believed that George W. Bush would not have been able to win in 2000 had Ralph Nader not run, since most Nader voters would then have voted for Gore.[1]

[1] In 1992, vote shares were 43.0% for Clinton, 37.5% for Bush and 18.9% for Perot. In 2000, vote shares were 47.9% for Bush, 48.4% for Gore and 2.7% for Nader. Despite getting fewer popular

Besides their potential to alter the election outcome, small parties are also important in the sense that they point to problems with simple plurality (winner-take-all) voting systems: Simple plurality systems lead to an underrepresentation of alternative political views (political views of people supporting minor parties) and create an incentive for supporters of minor parties to strategically misrepresent their true political preferences, which means that the election outcome is an imperfect measure of underlying true preferences.

Given that small parties generally have no chance of winning, it is rather surprising that people vote for them. One would normally expect that voters would not waste their votes for small parties, realizing that they can use their votes more effectively by voting for a potential winner. This is one of the reasons behind Duverger's Law, which states that simple plurality elections favor a two-party system and discourage third parties (Duverger, 1954). Hence, we should see weak party vote shares being zero in winner-take-all elections. However, this is not what we observe in the real world: Third parties have received an average of 5% of votes since 1920 in U.S. presidential elections. Small parties continue to exist and receive votes even though everyone knows that they have no chance of winning.

This paper addresses two main questions about small parties in simple plurality elections. The first question is why people vote for weak parties even when voters know they are going to lose anyway, and the second question is what factors explain the variation in weak party vote shares. Regarding the first question,

---

votes, Bush won the election by winning Florida with a difference of 500 votes, a state where Nader received 97,500 votes.

standard voting models predict zero vote share for parties with no chance of winning, thus fail to explain the empirical fact of positive vote shares for weak parties. Two prominent standard voting models are pivotal voter models and rule-utilitarian models.[2] Although these are two-party models, when the logic of these models is extended to three parties, they predict zero votes for weak third parties. According to the pivotal voter logic, people cannot be pivotal if they vote for a sure loser, and thus would never choose to vote for a weak third party. The same issue arises with rule-utilitarian models: there is no reason for a group to vote for a loser party if it cannot win.

A potential explanation for why people vote for parties with no chance of winning is that even when a voter votes for a front-runner, the probability that the voter is pivotal converges to zero as the number of voters approaches infinity. Expecting that they are not going to be pivotal anyway, voters might as well vote for their favorite party, or vote to express support for particular outcomes. This leads to the ideas of sincere and expressive voting:[3] If we assume that voters lack strategic

[2]Pivotal voter models assume that voters are individually rational and are motivated to vote by the chance that they might swing the election. Rule-utilitarian models assume that voters are ethically motivated to adopt the voting strategy that would maximize the aggregate utility of the community if everyone was to follow it. Prime examples are Ledyard (1984), and Palfrey and Rosenthal (1985) for pivotal voter models and Coate and Conlin (2004), and Feddersen and Sandroni (2006) for rule-utilitarian models.

[3]For examples of sincere voting models, see Palfrey (1984), Osborne and Slivinski (1996) and Callander (2005). For expressive voting, see Brennan and Buchanan (1984), Brennan and Lomasky (1997), Brennan and Hamlin (1998), Hamlin and Jennings (2011).

concerns, i.e. they do not take into account winning chances of each candidate, and they simply vote for the candidate that they like most (sincere voting) or vote to express support for outcomes associated with particular candidates (expressive voting), it is naturally the case that people who like weak parties will vote for them, so that weak parties get positive vote shares. However, sincere and expressive voting ideas fail to explain the empirical facts that voters often strategically desert weak parties and that weak party vote shares systematically covary with election-specific characteristics.[4,5]

In order to understand what is missing in standard voting models and the contribution of this paper, a discussion of instrumental and non-instrumental (intrinsic) benefits from voting is necessary. Instrumental benefits of voting are those that are related to the election outcome, such as raising the probability of having a more preferred candidate win the election over a less preferred candidate. Non-instrumental benefits or costs of voting are those that are independent of the election outcome, such as the benefit that arises from the satisfaction of citizenship duty or the effort cost of going to the polls.

In most of the previous literature, non-instrumental benefits of voting derive from the act of voting itself, and do not depend on which candidate the voter chooses. Candidate-specific benefits are only instrumental, i.e. they arise only from

[4] Using Japanese general election data, Kawai and Watanabe (2013) estimate that at least 64% of voters are strategic rather than sincere. They also find that the extent of strategic voting is higher in closer elections.

[5] This point is put forward by Mackie (2011) as well, in criticism of the expressive voting idea.

the possibility of having the candidate win the election. Voters take into account the non-instrumental benefits and costs when deciding on whether to vote or abstain, but if they decide to vote, there is no reason to vote for a candidate with no chance of winning, since the instrumental benefits of voting for a candidate with no chance of winning is zero.[6] In other words, standard models do not assign a differential non-instrumental benefit or cost associated with sincere vs. strategic voting.

In this paper, I introduce candidate-specific non-instrumental benefits as the explanation of positive vote shares for weak third parties with no chance of winning. Specifically, I suggest that there are non-instrumental benefits to sincere voting (instead of non-instrumental benefits of voting itself), i.e. there is an intrinsic benefit of voting for a candidate that best represents one's political views, which is not obtainable with strategic voting. Thus, for a voter whose favorite party is weak, the decision of whom to vote for involves a comparison of instrumental and non-instrumental gains. If he votes for his favorite party, he gets the non-instrumental benefit of sincere voting but forgoes the possibility of influencing the election outcome. If he deserts his favorite candidate and votes for a favorable contender, he gets the instrumental benefit of raising the probability of a more favorable outcome but gives up the non-instrumental benefit of sincere voting. I assume that non-instrumental benefits are heterogeneous across voters. For some voters, instrumen-

[6] An exception is the literature on expressive voting, where expressive benefits as well as instrumental benefits are candidate-specific. However, as discussed above, implications of expressive voting papers are at odds with empirical regularities. I discuss in the next footnote why they start from a similar point with this paper but end up with completely different implications.

tal benefits outweigh non-instrumental benefits, and for others, non-instrumental benefits outweigh instrumental benefits.[7] Thus, this model is able to generate positive weak party vote shares in equilibrium, an empirical fact that standard voting models cannot capture. It also generates partial strategic desertion of weak parties, an empirical fact that sincere voting models cannot capture.[8]

There naturally emerges the question of why there should be a non-instrumental benefit for voting sincerely. My explanations are as follows: First, people see elections as a chance to stand up and show support for their party. Supporting a party that describes you best has its own intrinsic benefit, even when you know that party is going to lose. Second, there is a value to the act of communicating true political preferences, arising from the desire to state one's true opinion and make it heard by others. Third, sincere voting saves face against people around you whose political views are similar to yours. Deserting a third party that you believe in for a mainstream likely winner can be frowned upon by others and likened to selling principles for material gain. A quote by British Labor Party politician Anne Begg nicely fits in: "Tactical voting is fine in theory and as an intellectual discussion in the drawing

[7] This is precisely where expressive voting papers fail. Despite empirical evidence on the contrary, these papers argue that expressive benefits necessarily dominate instrumental benefits for all voters in large elections, therefore instrumental benefits play no role in determining voter behavior, which fails to explain the existence of strategic voting.

[8] There are papers that include a mix of strategic and sincere voters, such as Kawai and Watanabe (2013) and Spenkuch (2013). However, being strategic or sincere is exogenously assigned to voters in these papers rather than being an optimal choice of voters, and hence these papers do not explain the mechanism by which voters become strategic or sincere.

room or living rooms around the country, but when you actually get to polling day and you have to vote against your principles, then it is much harder to do". I do not take this to mean that people never engage in strategic voting but rather infer that strategic voting is costly and people are only willing to do it when it is worth it. Therefore, the existence of non-instrumental costs to strategic voting (equivalent to non-instrumental benefits to sincere voting) is the defining assumption of this paper.

A closely related paper is Castanheira (2003), which asks why people vote for losers, essentially the same as the first question of this paper. The underlying idea is that people may vote for parties with no chance of winning, motivated by dynamic instrumental gains. For instance, an extreme leftist voter may vote for an extreme leftist party with no chance of winning, with the hope that this convinces mainstream parties to adopt more leftist policies in the next election. It is not clear, however, that voters only vote for instrumental reasons. Conventional political science wisdom suggests that excitement and emotions towards candidates play an important role in determining voter behavior. To the extent that sincere vs. strategic voting is a heart vs. mind decision and third party voting derives from non-instrumental gains, dynamic instrumentalism does not fully explain the phenomenon of voting for losers, thus a more general approach as in this paper is useful. Another important difference of this paper from Castanheira (2003) is that this paper generates a broader set of theoretical predictions (e.g. on the correlation between third party voting and turnout and differential turnout rates for major vs. minor party supporters) that are both consistent with existing empirical evidence and not that obvious ex ante.

I introduce heterogeneous intrinsic benefits to sincere voting into two otherwise standard models (pivotal voter and rule-utilitarian models) and tackle the second main question that this paper addresses: What are the factors that explain the variation in third party vote shares? Since votes for third parties are essentially sincere votes, this question equivalently asks when people are more likely to vote strategically vs. sincerely, i.e. when they are more likely to strategically desert third parties. In answering this question, I propose a framework with non-instrumental benefits to sincere voting, in which the strategic voting choice is endogenous and is affected by election-specific factors. Therefore, it is possible that the same voter with the same preferences may vote strategically in one election and sincerely in another election. I use a static framework with two strong parties and one weak party to investigate the behavior of weak party supporters. Sincere voting generates intrinsic utility regardless of the election outcome whereas strategic voting yields the potential benefit of changing the outcome of the election. Thus, given the costs and benefits of strategic and sincere voting, weak party supporters optimally choose whether to vote strategically or sincerely.

As one of the main results, this paper endogenously generates the stylized fact that the extent of strategic voting is higher in closer elections. This is an empirical result found by Spenkuch (2013) and Kawai and Watanabe (2013). Those papers add sincere voters to the strategic voting model of Myerson and Weber (1993) by assuming that voters are either sincere or tactical and that their types are assigned by nature. They both estimate the extent of strategic voting, using German and Japanese election data respectively. Both papers find empirically that strategic voting is more widespread in closer elections. However, these papers do not explain how voters become strategic, since their models start with the assumption that some voters are born strategic and others are born sincere. Conversely, the endogenous strategic voting framework I build in this paper explains why more voters choose to become strategic in closer elections.

Another relevant paper is Cox (1994), which also documents the relationship between strategic voting and closeness of the election by building a model of strategic voting to explain voting behavior in multimember districts, where $m > 1$ members from each district are elected. He finds a negative relationship between closeness (the difference of votes between m-th and $m + 1$-th candidate) and excess votes for leading candidates (sum of the difference of votes between first $m - 1$ candidates and the m-th candidate). Thus, he establishes that vote wasting on leading candidates in multimember districts (which could be interpreted as similar to vote wasting on loser candidates) is decreasing in closeness of the election.

The main results of this paper are as follows: For both pivotal voter and rule-utilitarian models, some fraction of third party supporters choose to desert their favorite party and vote strategically, while others stick to their favorite party and vote sincerely. Third party vote shares are lower in closer elections and when the stakes of the election are higher.[9] Pivotal voter and rule-utilitarian models have different implications about electorate size: Third party vote shares are higher in

[9]Stakes of the election represent the utility difference between major parties for third party supporters. Possible interpretations include political polarization, outrage against an incumbent, importance of the election and differentiated policy positions by major parties.

larger electorates for the pivotal voter model whereas the rule-utilitarian model does not yield clear-cut predictions on the effect of electorate size on third party vote shares. On the issue of turnout, this paper predicts that turnout rate is higher for supporters of viable parties than for supporters of third parties, and that variations in the competitiveness of the election create a negative correlation between third party vote shares and turnout whereas variations in the popularity of the third party candidate create a positive correlation between third party vote shares and turnout.

This paper contributes to the existing literature in the following ways: It takes two standard models (pivotal voter and rule-utilitarian models of costly voting), extends these models to three parties and introduces costly strategic voting into both models. By doing so, this paper generates predictions that are consistent with existing empirical evidence, which include the results by Kawai and Watanabe (2013) and Spenkuch (2013) that the extent of strategic voting is higher in closer elections, Bensel and Sanders (1979) that minor party supporters are more likely to abstain than major party supporters, Burden (2005) that strategic voting and turnout are positively correlated in nationally competitive elections whereas strategic voting and turnout are negatively correlated in nationally non-competitive elections. The main mechanism added by this paper (intrinsic benefits of sincere voting) is crucial in generating these results, since standard models without this mechanism fail to explain these empirical facts, as I discuss later in the paper.

Furthermore, to the best of my knowledge, this is the first paper in which strategic voting is endogenous (a choice) rather than exogenous (a type).[10] This paper also compares and contrasts the implications of pivotal voter and rule-utilitarian models with costly strategic voting. Finally, this paper presents new empirical evidence on third party voting in U.S. presidential elections from 1920 to 2012.

The rest of the paper is organized as follows: Section 1.2 presents the pivotal voter model with costly strategic voting and endogenous turnout. Section 1.3 presents the ethical voter (rule-utilitarian) model with costly strategic voting and exogenous turnout. Section 1.4 discusses existing relevant empirical evidence and presents new empirical evidence from U.S. presidential elections. Section 1.5 concludes.

1.2 Pivotal Voter Model

I build a pivotal voter model in this section to investigate the issues of voting for third parties and strategic voting. I adapt the Palfrey and Rosenthal (1985)

[10] Most of the previous literature starts with the assumption that all voters are strategic (they vote for the candidate who maximizes the expected instrumental utility of voting) or that all voters are sincere (they vote for the candidate they like most, regardless of the candidate's chances of winning). Papers that combine strategic and sincere voters such as Spenkuch (2013) and Kawai and Watanabe (2013) assume that nature exogenously assigns some voters to be strategic and others to be sincere. Note that by strategic voting, I here mean voting for the candidate that maximizes expected instrumental utility of voting, which does not necessarily mean voting for a candidate other than one's favorite. Kawai and Watanabe (2013) calls the latter 'misaligned voting' to distinguish the two and misaligned voting is endogenous in many previous papers.

model to the three candidate endogenous turnout case in order to study voter behavior in an election with two strong candidates and one weak candidate.

Below is my version of the calculus of voting equation:[11]

$$R_{ij} = \frac{1}{2} \sum_{k \neq j} p_{ijk}(B_{ij} - B_{ik}) + D_{ij} - C_i$$

where R_{ij} is the expected payoff for voter i if he votes for candidate j,[12] p_{ijk} is the probability of being pivotal,[13] B_{ij} is the instrumental utility of having candidate j in office for voter i and D_{ij} is the intrinsic utility that voter i gets by voting for candidate j and C_i is the cost of voting.

I preserve the endogenous turnout aspect of Palfrey and Rosenthal (1985), where voting is costly and individuals decide whether to vote or not based on how costs and benefits of voting compare.[14] The departure from the conventional calculus of voting equation is the D term. In the standard calculus of voting framework, D is the benefit that the voter derives from the act of voting itself, due to the sense of

[11]See Section 1.2.5 for the standard calculus of voting equation and how its implications differ from the implications of my version.

[12]This can be thought of as the expected payoff of voting relative to not voting, where the payoff in the case of not voting is normalized to zero.

[13]Ties are resolved with a fair coin toss. Thus pivot events occur when candidate j is either tied or one vote behind candidate k, not counting voter i's vote. In case of pivot events, voter i receives an expected utility gain of either $B_{ij} - \frac{1}{2}(B_{ij} + B_{ik})$ or $\frac{1}{2}(B_{ij} + B_{ik}) - B_{ik}$ by voting for candidate j, both of which equal $\frac{1}{2}(B_{ij} - B_{ik})$.

[14]An earlier version of the model with exogenous turnout (where each voter is assumed to vote with a constant probability π) yields the same qualitative predictions on third party voting. The endogenous turnout model is richer since it yields additional predictions on voter participation.

citizen duty. A voter gets this benefit if he votes, no matter which candidate he votes for. Conversely, I argue that the intrinsic utility term derives from backing the party that best represents your views, in which case D depends on which candidate the voter selects. A voter gets this utility only when he votes for his favorite candidate, not when he votes strategically for another candidate that has a better chance of winning the election. Thus, I specify the structure of the D term as follows:[15]

$$D_{ij} = \begin{cases} D_i & \text{if } B_{ij} \geq B_{ik} \text{ for any candidate k } (j \text{ is the favorite candidate of } i) \\ 0 & \text{otherwise} \end{cases}$$

The value of voting for voter i is the maximum utility that voter i can get from voting:

$$V_i = \max_j R_{ij}$$

Voter i votes for the utility-maximizing candidate if $V_i > 0$ and abstains if $V_i \leq 0$. Suppose there are three candidates, N "moderate" voters and N^E "extreme partisan" voters. Each voter is in one of three groups: T_1, T_2 and T_3. The favorite candidate of a voter in group T_j is j. The number of moderate voters in T_j is N_j and the number of extreme partisans in T_j is N_j^E, where $N_1 + N_2 + N_3 = N$ and $N_1^E + N_2^E + N_3^E = N^E$. Each voter has four options: vote for candidate 1, vote for candidate 2, vote for candidate 3 or abstain. However, extreme partisans always turn

[15] I could allow D_{ij} to be positive for more than one candidate, which would not change the results qualitatively for the three-candidate setup considered in this paper, as long as D_{ij} is highest for the voter's favorite candidate.

out to vote for their favorite candidate, i.e. extreme partisans in T_j always vote for j and never abstain.[16] On the other hand, moderate voters could potentially abstain, vote for their favorite candidate, or vote for their second favorite candidate. The candidate with the highest number of votes wins the election and ties are resolved with a fair coin toss.

For all voters, voter preferences over candidates are specified as $B_{ij} = z - |x_j - x_i|$ where x_i represents the location of the favorite candidate of voter i and x_j represents the location of candidate j along the political spectrum. Suppose that Candidate 2 is located between Candidates 1 and 3. Let us normalize the distance between Candidate 1 and Candidate 3 as the utility of having one's favorite party win the election, i.e. z units. Suppose the distance between Candidate 1 and Candidate 2 is y units, which implies that the distance between Candidate 2 and Candidate 3 is $z - y$ where $z > y$. Thus, we have $B_{i1} = z, B_{i2} = z - y$, and $B_{i3} = 0$ for voters in T_1; $B_{i1} = z - y, B_{i2} = z$, and $B_{i3} = y$ for voters in T_2; and $B_{i1} = 0, B_{i2} = y$, and $B_{i3} = z$ for voters in T_3. Thus, for example, if a voter i in T_1 votes for candidate 1 and this vote becomes pivotal in making Candidate 1 win the election over Candidate 2, the utility gain to the voter is $B_{i1} - B_{i2} = y$ units. Also, notice that the second choice of both voters in T_1 and T_3 is Candidate 2, while the second choice of voters in T_2 could be Candidate 1 or Candidate 3 (depending on the values of y and z). When a voter decides to vote strategically, he votes for his second choice (voters never vote for their last choice).

[16] Assuming $D_i = \infty$ for extreme partisans is consistent with this behavior.

Now that we have restricted the analysis to the three candidate case and specified the structure of B_{ij}'s, we can use the calculus of voting equation to write the expected payoffs of voting (R_{ij}'s) for each candidate for voters in each group, which I provide in Appendix A.1.

I assume that the sincere voting benefit D_i and the cost of voting C_i are stochastic and private information. That is, a voter knows his actual D_i and C_i and the distribution of D_i and C_i for others. On the other hand, the number of voters in each group (N_1, N_2 and N_3) and preferences of voters over candidates (B_{ij}'s) are commonly known by all voters.

Let us assume that the sincere voting benefit is independently and identically distributed across all voters in the electorate, and its distribution is uniform with lower bound 0 and upper bound $\bar{D}$, where $\bar{D} > 0$. Similarly, the cost of voting is independently and identically distributed across all voters according to a uniform distribution lower bound 0 and upper bound $\bar{C}$. Hence, we have:

$$D_i \sim U(0, \bar{D}) \ \forall i$$

$$C_i \sim U(0, \bar{C}) \ \forall i$$

Since this paper is about third party voting, I focus on a setting with two strong front-runners and one weak third party that is sure to lose. I make party 1 and party 2 the strong parties and party 3 the weak party. Since party strength is determined by the number of voters that have the party as their first choice, I concentrate on the case where the number of voters in T_3 is significantly smaller than the number of voters in T_1 or T_2. Specifically, I concentrate on cases where the number of extreme partisans in T_1 and T_2 are both greater than total number of voters in T_3, i.e. $N_1^E > N_3^E + N_3$ and $N_2^E > N_3^E + N_3$. This makes it impossible for party 3 to win the election unless moderate voters in T_2 decide to vote for candidate 3. Even if all voters in T_3 vote for party 3, it would not be enough to pass the extreme partisan votes of party 1 and party 2.

Theoretically, the above assumption about group sizes does not prevent party 3 from being a front-runner since this model has multiple equilibria. If voters in T_2 prefer party 3 over party 1 (if $y > z/2$), there may exist an equilibrium where the majority of voters in T_2 vote for party 3, thus making party 3 a front-runner. However, I focus on the equilibrium where voters in T_2 stick to party 2 (their first choice) rather than voting for party 3 (their second choice). This equilibrium makes more sense since voters in T_2 can eliminate the possibility of a party 3 win and can make their favorite party a front-runner by simply voting for their favorite party. This is also the unique equilibrium for $y < z/2$.

When $N_1^E > N_3^E + N_3$, $N_2^E > N_3^E + N_3$, and moderate voters in T_1 and T_2 vote for their favorite party, the probability of a party 3 win becomes zero, so that pivot probabilities involving party 3 ($p_{13}, p_{23}, p_{31}, p_{32}$) are all zero. Using R_{ij}'s provided in Appendix A.1, this implies $R_{i2} < 0$ and $R_{i3} < 0$ for voters in T_1; $R_{i1} < 0$ and $R_{i3} < 0$ for voters in T_2; $R_{i1} < 0$ for voters in T_3. This means that the choice problem for voters in T_1 and T_2 is whether to vote or abstain (since it is optimal for them to vote for 1 and 2 respectively if they vote), whereas voters in T_3 decide on both whether to vote and for whom to vote.

R_{ij}'s provided in Appendix A.1 imply the following: A voter in T_1 votes (for party 1) if $p_{i12}y + D_i > C_i$ and abstains otherwise. A voter in T_2 votes (for party 2) if $p_{i21}y + D_i > C_i$ and abstains otherwise. A voter in T_3 votes sincerely for his favorite party 3 if $D_i > C_i$ and $D_i > p_{i21}y$, votes strategically for his second choice party 2 if $D_i < p_{i21}y$ and $p_{i21}y > C_i$, abstains if $D_i < C_i$ and $p_{i21}y < C_i$. Now we can define the equilibrium as follows:

Definition. *A Bayesian Nash Equilibrium is a set of thresholds $t^* = (t_1^*, t_2^*, t_3^*)$ such that voters in T_1 vote for candidate 1 if $C_i - D_i < t_1^*$ and abstain otherwise; voters in T_2 vote for candidate 2 if $C_i - D_i < t_2^*$ and abstain otherwise; voters in T_3 vote for candidate 3 if $D_i = max(D_i, C_i, t_3^*)$, vote for candidate 2 if $t_3^* = max(D_i, C_i, t_3^*)$ and abstain if $C_i = max(D_i, C_i, t_3^*)$.*

Let p_1^* denote p_{i12} for voters in T_1, p_2^* denote p_{i21} for voters in T_2, p_3^* denote p_{i21} for voters in T_3 in equilibrium.[17] Equilibrium thresholds are then given by $t_1^* = p_1^*y$, $t_2^* = p_2^*y$, $t_3^* = p_3^*y$. I will now state and prove existence of a PBE.

Proposition 1. *There exists an equilibrium $t^* = (t_1^*, t_2^*, t_3^*)$ with $t_1^* \in [0, y], t_2^* \in [0, y], t_3^* \in [0, y]$.*

Proof. See Appendix A.2.

I solve the model using computational methods. The basic idea is to start with initial guesses for p_1, p_2 and p_3, which then imply the thresholds adopted by voters in three groups by $t_1 = p_1y$, $t_2 = p_2y$, $t_3 = p_3y$. These thresholds imply

[17]Note that it is okay here to drop the i subscripts since pivot probabilities are the same for voters in the same group.

the probability that a voter votes in T_1, the probability that a voter votes in T_2, the probability that a voter in T_3 votes sincerely and the probability that a voter in T_3 votes strategically. These in turn imply probabilities of being pivotal p_1, p_2 and p_3. Equilibrium probabilities p_1^*, p_2^*, p_3^* are found where the initial guesses for p_1, p_2, p_3 are equal to the resulting probabilities of being pivotal. I provide in detail the computational algorithm that solves the model in Appendix A.3.

The baseline parameter values are $N_1 = 1000, N_2 = 950, N_3 = 100, y = 40, \bar{D} = 1, \bar{C} = 2$.[18] As defined earlier, N_k is the number of moderate voters in T_k for $k = 1, 2, 3$, y is the utility gain (utility loss for voters in T_1) of having party 2 win the election instead of party 1 for voters in T_2 and T_3, $\bar{D}$ is the upper bound of the sincere voting benefit distribution, $\bar{C}$ is the upper bound of the voting cost distribution.[19]

In the following, I present the results of experiments regarding the effects of a subset of parameters on endogenous variables. These results will tell us this model's predictions on the determinants of third party vote shares, the extent of

[18] As for the number of extreme partisans, as long as $N_1^E > N_3^E + N_3$ and $N_2^E > N_3^E + N_3$ are satisfied, all that matters is the difference in the number of extreme partisans in T_1 and T_2, i.e. $N_1^E - N_2^E$. I use $N_1^E = N_2^E$ or $N_1^E - N_2^E = 0$ for the baseline case.

[19] These parameters represent a relatively small election with electorate size around 2000 voters. The reason is that the computational cost is increasing exponentially with electorate size. The same results can be obtained for larger electorates (where pivot probabilities are lower) by using higher $y/\bar{D}$ and $y/\bar{C}$ ratios. The qualitative results I present in the body of the paper are reasonably general and robust to a wide range of parameter choices. I discuss results that are particularly sensitive to the parameter values in Appendix A.4.

strategic voting and voter turnout. Parameters that are not the subject of the specific experiment are kept at their baseline values stated above.

1.2.1 Indifference Towards Major Parties

The first set of results concerns the effect of the utility difference between the major parties, given by $|B_{i2} - B_{i1}|$, which is equal to y for all voters. As y rises, voters get a higher utility gain of having the preferred front-runner (party 1 for voters in T_1, party 2 for voters in T_2 and T_3) win the election over the other. Conversely, as y gets closer to zero, voters become indifferent towards major parties. I refer to y as the "stakes of the election" for all voters.

Panels a and b of Fig. 1.1 plot the expected fraction of strategic and sincere voters in T_3, and vote shares of each candidate as functions of y. Among third party supporters (voters in T_3), the extent of strategic voting is increasing and the extent of sincere voting is decreasing in y. Consequently, the third party vote share (vote share of party 3) is decreasing in y.

This is an intuitive result, saying that as major parties make less effort to distinguish their policies from each other, voters at the ends of the political spectrum will be less inclined to vote for them. More voters will therefore opt to vote for parties that better represent their preferences even when those parties are weak. This can also be interpreted as voters penalizing center parties for not creating enough appeal, for instance an extreme leftist voter penalizing a center-left party for not adopting sufficiently leftist policies.

On the other hand, this result says that a higher extent of strategic desertion of small parties can be expected when one of the major parties irritates certain groups of the electorate and causes intense outrage. This can be represented by a fall in B_{i1} for voters in T_3, and hence a rise in y, in which case the model implies a higher fraction of strategic voters and a lower vote share for small parties. When one of the major parties irritates a certain fraction of the electorate, one can expect irritated voters to vote for the strongest challenger to that major party instead of their weaker favorite candidate, and hence exhibit widespread strategic voting, which causes the vote share of third parties to fall.

Panel c of Fig. 1.1 plots the participation rates for different groups and for the whole electorate as functions of y. Participation rates for all groups as well as the participation rate for the electorate as a whole are increasing in y. The reasons that induce higher turnout by voters in T_1 and T_2 are the same as those that induce a higher extent of strategic voting by voters in T_3. As voters in T_1 and T_2 differentiate more between the front-runners, expected payoff of voting rises and they get more inclined to turn out to vote rather than abstain.

Panel c of Fig. 1.1 also demonstrates that the participation rate for party 3 supporters are lower than participation rates of party 1 and party 2 supporters. This is a general result that will hold true for the other exercises as well (as long as distributions of the voting cost and the sincere voting benefit are the same across groups), and is one of the core implications of this model with costly turnout and costly strategic voting: A strong party supporter gets both instrumental and non-instrumental benefits by voting for his most preferred party whereas a third party supporter only gets one of these benefits since he has to forgo the expected benefit that arises from the probability of a pivot event if he votes sincerely or he has to forgo the sincere voting benefit if he were to vote strategically for a stronger party. Therefore, the total benefit that arises from the act of voting is higher for major party supporters than minor party supporters, which implies that a greater proportion of major party supporters will participate than minor party supporters. This theoretical prediction is consistent with the empirical findings of Bensel and Sanders (1979). Using data from 1968 U.S. presidential elections, they find that the highest percentage of non-voting is found on those who favor the minor party in their states, which means that minor party supporters are more likely to abstain than major party supporters.

Panel d of Fig. 1.1 plots equilibrium probability of being pivotal for voters in T_1, T_2 and T_3 as functions of y. The movements of pivot probabilities reflect two effects. As y rises, increased number of voters due to increased turnout reduces the probability that any vote is pivotal, whereas a higher extent of strategic voting increases pivot probabilities at first and reduces them eventually through its impact on the closeness of the election. These generate an overall downward trend accompanied by a spike in pivot probabilities in the middle where increased strategic voting by third party supporters creates a strongly positive marginal impact on the closeness of the election between front-runners.

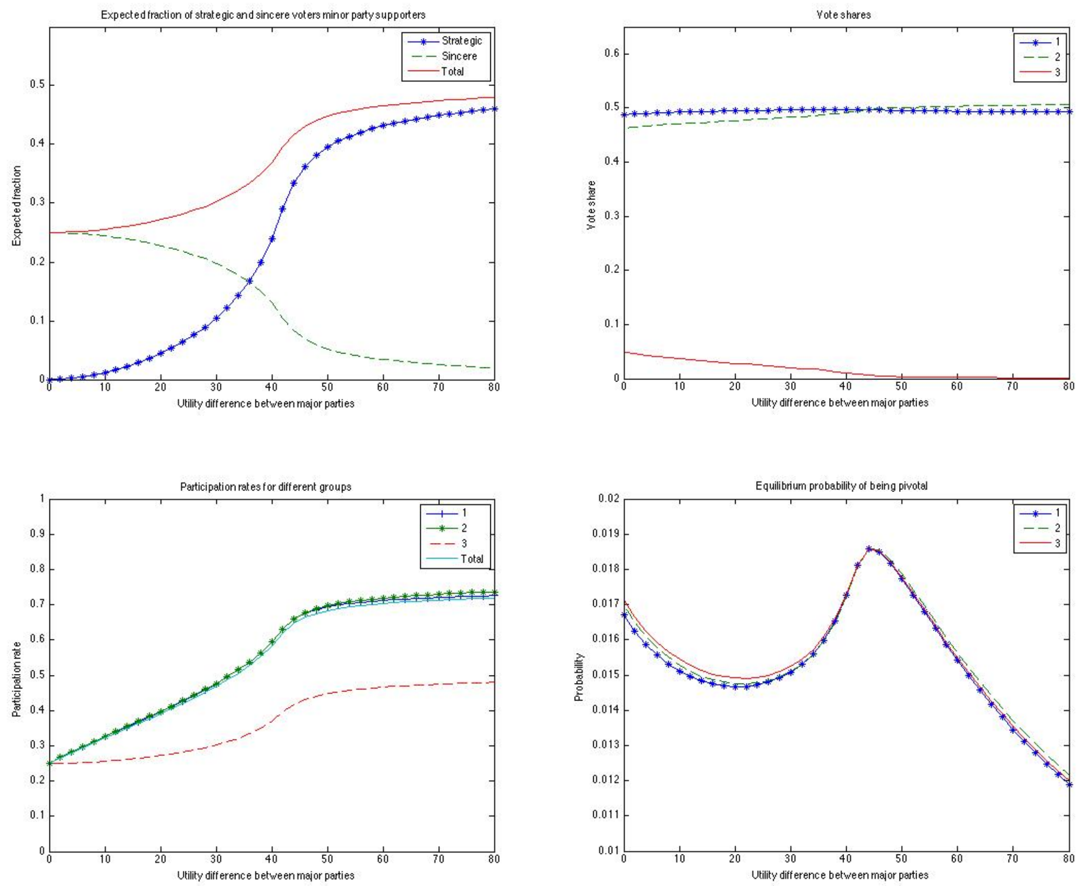

Figure 1.1: Third Party Voting vs. Utility Difference Between Major Parties: (a) sincere and strategic voting, (b) vote shares, (c) participation rates, (d) pivot probabilities

1.2.2 Closeness of the Election

The second set of results relates to the effect of the difference in strength between major parties, defined to be $x = N_1 - N_2$, i.e. the difference between the number of supporters for party 1 and party 2. As x gets closer to zero, the election is more likely to be close. I emphasize that this is a rough but not an exact measure of closeness because taking into account the strategic voters in T_3, a level of x that is small and positive can be expected to generate a closer election than $x = 0$.

Panel a of Fig. 1.2 plots the expected fractions of strategic and sincere voters in T_3 as functions of x. When x is negative and sufficiently large, meaning that N_2 is sufficiently greater than N_1, strategic voting almost disappears for voters in T_3, since the number of votes coming from T_2 is already high enough in expectation for party 2 to beat party 1 easily, so strategic behavior by party 3 supporters is not necessary. Strategic behavior also disappears when x is positive and sufficiently large, since party 1 is expected to win the election easily even if all voters in T_3 vote for party 2. Hence, the model implies that strategic behavior will vanish when one side is perceived to be sufficiently stronger than the other. When the election is sufficiently one-sided, it is not worthwhile for third party supporters to forgo their non-instrumental sincere voting benefit in the hopes of helping a more favorable contender win the election.

The number of strategic voters rises as x gets closer to zero. Interestingly, strategic voting is most wide-spread and the third party vote share is lowest when x is small but positive instead of zero. The reason is that accounting for the strategic voters in T_3, the election is expected to be closer (as suggested by the graph of the equilibrium probabilities of being pivotal) when x is slightly higher than zero and below some threshold. Hence, the result that strategic voting will be more widespread in closer elections withstands.

The extent of strategic voting being higher and third party vote shares being lower (panel b of Fig. 1.2) in closer elections result from the optimal response of third party supporters to the changes in pivot probabilities (panel d of Fig. 1.2). As the election gets closer between front-runners, probability of affecting the election outcome by voting for a viable party increases. Responding to that, more third party supporters decide to vote strategically with the hope of affecting the election outcome.

Panel c of Fig. 1.2 plots participation rates as functions of x. Voter turnout rises as the election gets closer between the front-runners. Closer races between front-runners create spikes in pivot probabilities, which induce both strategic desertion of third parties by their supporters and higher participation by the electorate as a whole.

The strategic voting result contributes to the literature on strategic voting by providing a theoretical explanation for the empirical findings of Kawai and Watanabe (2013) and Spenkuch (2013) that the extent of strategic voting is higher in closer elections. These papers do not explain this finding theoretically, since voters being strategic vs. sincere is exogenously determined in those models. The turnout result is also consistent with the ethical voter model of Coate and Conlin (2004), which establishes a positive correlation between turnout and closeness of elections.

Fig. 1.2 also demonstrates a more general (and perhaps obvious ex post but less obvious ex ante) implication of this paper on the correlation between turnout and third party vote shares. In response to variations in pivot probabilities, the model predicts a negative correlation between third party vote shares and turnout. When there are intrinsic benefits of sincere voting (equivalent to intrinsic costs of strategic voting), voting and strategic voting have costs associated with them, and it is more worthwhile to pay these costs and engage in both of them in elections where the probability of changing the election outcome or the stakes of the election are high. This causes voter turnout and the extent of strategic voting to move in the same direction, which generates a positive correlation between strategic voting and turnout, hence a negative correlation between third party vote shares and turnout.

This prediction of a positive correlation between strategic voting and turnout is consistent with the empirical evidence by Burden (2005). He finds that strategic desertion of third party candidates and voter turnout are positively correlated in the 2000 U.S. presidential election, where the Electoral College was very competitive and therefore state-level closeness of the election between front-runners (variations in p) would be expected to mainly generate the correlation between third party voting and turnout.

1.2.3 Size of the Electorate

Another factor affecting strategic voting is the size of the electorate. The size of the electorate is important for a pivotal voter model because the probability that

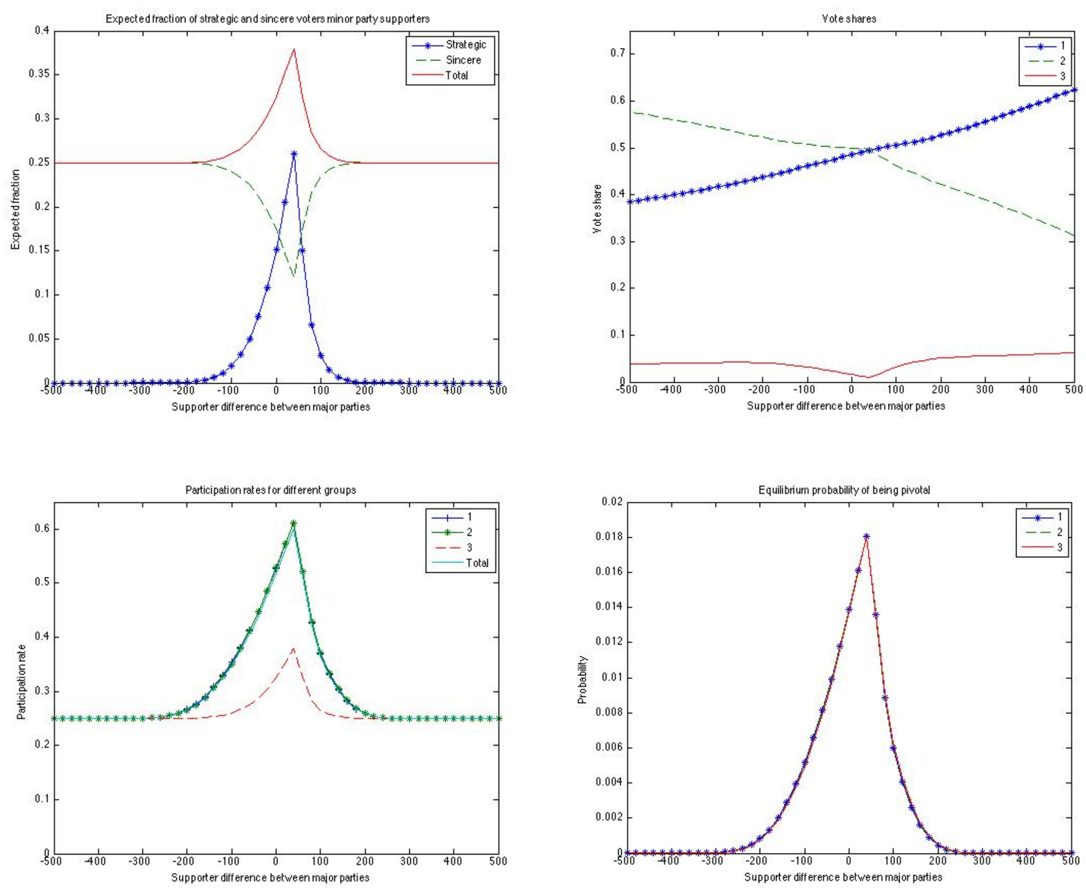

Figure 1.2: Third Party Voting vs. Difference in the Number of Supporters Between Major Parties: (a) sincere and strategic voting, (b) vote shares, (c) participation rates, (d) pivot probabilities

the election will be determined by a single vote decreases as the electorate gets larger.

To examine the impact of electorate size on strategic voting, I fix the ratios of the number of supporters in each group to the total number of voters in the electorate and multiply these ratios by different scalars, in order to study the implications of the size of the electorate while controlling for power differences across parties.[20]

Panels a and b of Fig. 1.3 plot the expected fractions of strategic and sincere voters, and vote shares of each party as a function of the electorate size. The third party vote share is increasing and the expected fraction of strategic voters is decreasing in the electorate size. These are both optimal responses to the decreasing probability of a pivot event (Panel d of Fig. 1.3) as the electorate gets larger. As the electorate gets larger, strategic voting for the more favorable contender loses its appeal since a pivot event is increasingly unlikely. Observing this, a higher fraction of third party supporters decide to vote for their favorite party instead of a less preferred party with higher chances of winning.

The result that strategic voting decreases in larger electorates is potentially useful in evaluating the success of the pivotal voter model. This result will not hold for the group rule-utilitarian model of the next section, which provides a way to assess the relative performance of these two models in terms of explaining the data.

[20]For this exercise, ratios are set to be $\frac{N_1}{N_1+N_2+N_3} = \frac{20}{41}$, $\frac{N_2}{N_1+N_2+N_3} = \frac{19}{41}$, $\frac{N_3}{N_1+N_2+N_3} = \frac{2}{41}$ and scalars are set to range from 1025 to 3075, which means that the electorate size $(N_1 + N_2 + N_3)$ also ranges from 1025 to 3075. The benchmark parameters are achieved when the scalar is 2050.

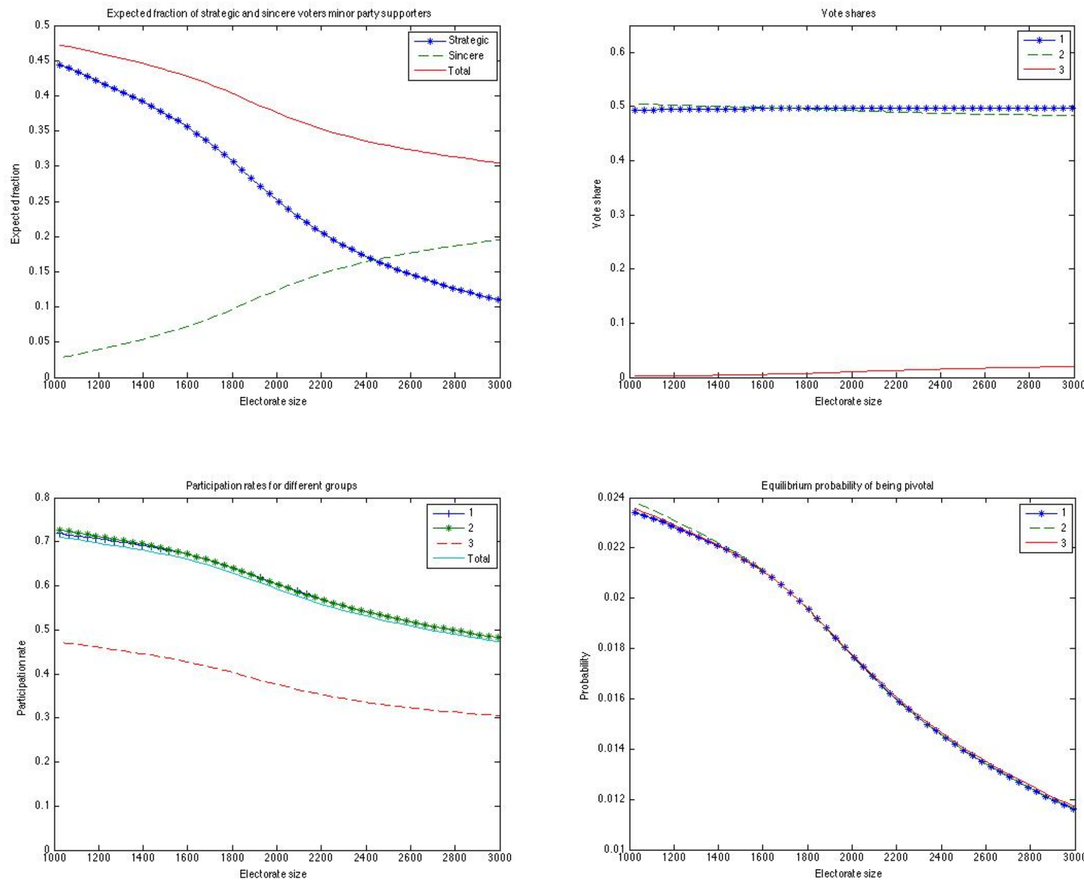

Figure 1.3: Third Party Voting vs. Electorate Size: (a) sincere and strategic voting, (b) vote shares, (c) participation rates, (d) pivot probabilities

Panel c of Fig. 1.3 plots participation rates as functions of the electorate size. Voter participation rates for all groups are decreasing in the electorate size, since the expected payoff of voting decreases in larger electorates due to pivot events being less likely. Voter turnout being lower in larger electorates is another theoretical prediction that is potentially useful in testing how successful the pivotal voter model is in terms of explaining the data.

1.2.4 Popularity of the Third Party Candidate

Popularity of the third party candidate is obviously one of the main determinants of the third party vote share. A popular third party candidate would be described as the one that yields higher benefits (both instrumental and non-instrumental) to his supporters. Since instrumental benefits associated with the third party candidate does not matter (he can never win the election), I look at the impact of non-instrumental benefits associated with the third party candidate in this experiment. To do so, I allow $\bar{D}$ to be different across groups for this exercise and examine the impact of changes in $\bar{D}_3$, that is, the upper bound of the sincere voting benefit distribution for voters in T_3.[21] To isolate the impact of $\bar{D}_3$ alone, I pick a non-close election with $x = -300$, where variations in pivot probabilities due to changes in $\bar{D}_3$ are small. All other parameters are set at their benchmark values including $\bar{D}_1$ and $\bar{D}_2$.

Panels a and b of Fig. 1.4 plot the expected fractions of strategic and sincere voters in T_3 as well as vote shares for each party as functions of $\bar{D}_3$. The extent of sincere voting and the vote share of the third party is increasing in popularity of the third party candidate $\bar{D}_3$. Moreover, panel c of Fig. 1.4 shows that a popular third party candidate significantly increases turnout by third party supporters (voters in T_3), which translates into a modest increase in overall turnout since third party supporters are the minority. Taken together, these predictions imply that popular

[21] Since sincere voting benefit is uniformly distributed, increasing the upper bound implies increasing the mean of sincere voting benefit as well.

minor party candidates both get higher vote shares and induce extra turnout by voters who would have abstained otherwise. In contrast to the previous results, the model therefore implies a positive correlation between third party vote shares and turnout in response to changes in third party candidate popularity.

These predictions are consistent with the empirical findings of Lacy and Burden (1999) and Burden (2005). Using data from 1992 U.S. presidential elections, Lacy and Burden (1999) find that the candidacy of Ross Perot (who obtained 19% of the nation-wide vote share in that election) increased voter turnout by around three percentage points. On the other hand, Burden (2005) detects a negative correlation between strategic desertion of third party candidates and turnout in 1992 and 1996 U.S. presidential elections, where the Electoral College was not competitive and therefore third party candidate popularity effects are expected to mainly generate the correlation between third party voting and turnout.

Appendix A.4 presents another experiment with the pivotal voter model on the effects of the minority group size (N_3). This experiment does not yield particularly strong comparative statics results.

1.2.5 Comparison to Standard Models

To summarize the results of the pivotal voter model: The third party receives a positive vote share in equilibrium, the third party vote share is lower in closer elections and when the utility difference between major parties for third party supporters is higher, and the third party vote share is higher in larger electorates. On

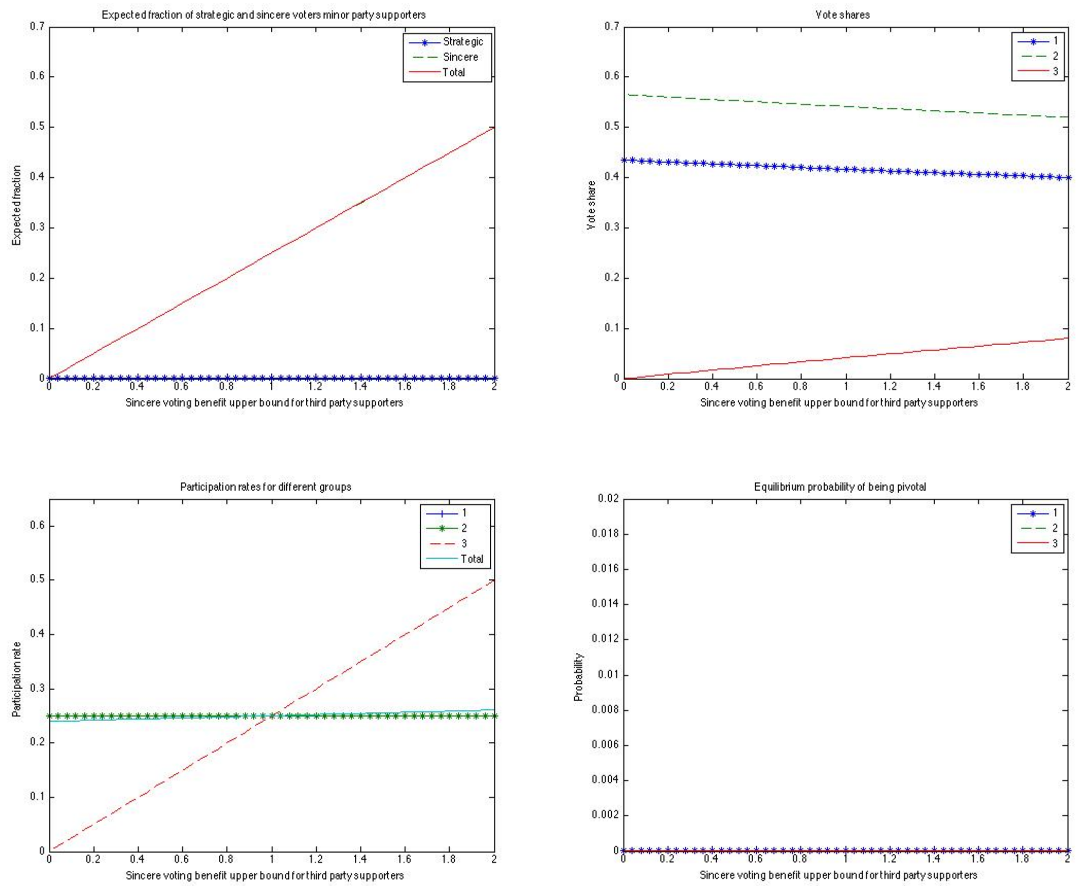

Figure 1.4: Third Party Voting vs. Third Party Candidate Popularity: (a) sincere and strategic voting, (b) vote shares, (c) participation rates, (d) pivot probabilities

the issue of turnout, the model predicts that minor party supporters have a higher abstention rate than major party supporters, and that variations in pivot probabilities generate a negative correlation between third party voting and turnout whereas variations in third party candidate popularity generates a positive correlation between third party voting and turnout.

All of these results depend on costly strategic voting. To see why, suppose that, as in the standard pivotal voter model, strategic voting is not costly in that the intrinsic benefit of voting is not candidate-specific (D_i instead of D_{ij}). In this case, the calculus of voting equation would be as follows:

$$R_{ij} = \frac{1}{2} \sum_{k \neq j} p_{ijk}(B_{ij} - B_{ik}) + D_i - C_i$$

In this case candidate choice has no effect on the intrinsic benefit D term because voter i will receive D_i no matter which candidate he votes for. Thus, candidate choice is solely determined by the comparison of instrumental benefits of each candidate (the first term of R_{ij}). Since voting for a candidate with no chance of winning offers no instrumental benefits, it is never optimal to vote for a loser candidate, even when the candidate is the one you like most. Thus, the standard pivotal voter model predicts that weak third parties always get zero votes, independently of model parameters such as closeness of election, utility difference between major parties and electorate size.

This also breaks the correlation between third party voting and turnout, since turnout is responsive to above model parameters in the standard model whereas the third party vote share is not. Moreover, abstention rate for minor party supporters would be the same as major party supporters in the standard model as long as they have the same instrumental utility difference between major parties as major party supporters.

Thus, the simple addition of costly strategic voting enables this model to generate many empirical facts that the standard pivotal voter model is unable to explain: positive third party vote shares, a higher extent of strategic voting in closer elections, a stronger tendency for minor party supporters to abstain compared to major party supporters, a negative correlation between third party voting and turnout when competitiveness of the race is the main factor and a positive correlation between third party voting and turnout when candidate-specific benefits are more important. The model also generates other main predictions (third party vote shares decreasing in the utility difference between major parties and increasing in the electorate size), which are suitable for empirical testing.

1.2.6 Criticism of the Pivotal Voter Model

The well-known criticism of the standard pivotal voter model is that since pivot probabilities go to zero as the number of voters rise, the only people who vote in large elections will be those whose intrinsic benefit of voting D_i is higher than their cost of voting C_i, i.e. those with negative net cost of voting $c_i = C_i - D_i$. Those with positive net cost of voting will abstain, understanding that they will not be pivotal anyway. Thus, the standard pivotal voter model can predict positive turnout through negative net voting costs but cannot predict the effect of election-related factors such as closeness between front-runners on turnout.

In evaluating this criticism, it is first necessary to make clear that the probability of being pivotal approaches zero but is never actually zero in practice, since the number of voters is always finite even for the largest electorate. Therefore, a more accurate statement of the criticism is that since the probability of being pivotal is very low in large elections, the pivotal voter model requires that instrumental utility differences be very high compared to voting costs and non-instrumental benefits, i.e. it requires very high y/C or y/D ratios, in order for closeness to have a non-trivial marginal impact on turnout.[22]

Before stating the arguments in defense of the pivotal voter model, I will first argue that this criticism is not of first order importance for this paper. My point is as follows: The standard pivotal voter model requires very high y/c ratios to predict that turnout is increasing in closeness of the race, but even with a very high y/c, it cannot explain the third party voting patterns. Hence, there are two distinct problems associated with the standard pivotal voter model: Requiring y/C and y/D to be very high to create quantitatively significant comparative statics, and failing to match the empirical facts regarding third party voting. I address the latter problem with this paper, which I show can be fixed by adding the notion of

[22] To give a sense of the orders of magnitude, for an election with 4 million voters (around the median voting age population across U.S. states) divided equally between opposite parties where each voter votes with 60% probability, the probability that a voter is pivotal is 0.0004, which implies that y/c ratio needs to exceed 2500 for a voter to vote.

non-instrumental benefits to sincere voting, whereas I relegate the former problem to other papers.

In case the reader still thinks I should address the former problem since I am using the standard pivotal voter model as benchmark, I will state the two main arguments in defense of the pivotal voter model. One is the notion of other-regarding social preferences, according to which instrumental benefits include not only private but also social benefits since people care about the well-being of others. Jankowski (2002) resembles a costly vote to "a lottery ticket to help the poor". Edlin, Gelman and Kaplan (2007) make the point that when social benefits at stake are large, expected benefits of voting for an individual with social preferences can be significant. Myatt (2015) shows that even very mild social preferences (where individuals put a very small weight on the utility of others compared to themselves) can generate very high y/c ratios.

The second argument is that voters may be motivated to vote by a pivot event because they may be overweighting the probability of a pivot event, as in Kahneman and Tversky (1992). When voters overweight pivot probabilities, variations in pivot probabilities due to changes in the closeness of the election can have a non-trivial impact on voter turnout without requiring a very high y/c ratio.

In the next section, I introduce the same mechanism (non-instrumental benefits to sincere voting) to another prominent voting model, the group-rule utilitarian model. I will show that most of the qualitative results of the pivotal voter model regarding third party voting extend to the group-rule utilitarian model, indicating that it is not the model choice but the notion of costly strategic voting that produces plausible results on third party voting.

1.3 Group Rule-Utilitarian Model

In this section, I build a group rule-utilitarian (ethical voter) model with costly strategic voting. In group-rule utilitarian models, each citizen adopts the strategy that would maximize the utility of the group if everyone in the group were to follow it. Citizens are ethically but not instrumentally motivated to do so. Since an individual ethical voter assumes that the strategy adopted by him will also be followed by other ethical voters in his group, he does not feel atomistic in terms of changing the election outcome when adopting a strategy. Therefore, rule-utilitarian models generate substantial turnout even in very large electorates without requiring very small voting costs.

The model is based on the ethical voter models of Coate and Conlin (2004) and Feddersen and Sandroni (2006). These are costly voting models of turnout whereas I study a costly strategic voting model. My setup is as similar as possible to the pivotal voter model to be able to make a fair comparison of models. Nevertheless, there are differences that I describe below.

The main difference is the exogenous turnout in the ethical voter model vs. costly voting and endogenous turnout in the pivotal voter model. For the following ethical voter model, instead of voters deciding on whether to turn out to vote given their voting costs, I introduce exogenous randomness on total turnout rates. Specifically, I assume that the turnout rate is $\tilde{q_1}$ for group 1 and $\tilde{q_2}$ for group 2, where $\tilde{q_1} \sim U(b,1)$ and $\tilde{q_2} \sim U(b,1)$. The turnout rate is deterministic for group 3 and is equal to $\frac{b+1}{2}$, so that turnout rates are the same for all three groups on average.[23]

As in the pivotal voter model, there are three parties and voters can be separated into three groups according to their favorite party. Voter preferences over parties are the same as in the pivotal voter model. Voters in group 1 prefer party 2 over party 3, and voters in group 3 prefer party 2 over party 1 (with a utility difference of y units). Voter preferences over parties are public information.

As in the pivotal voter model, each voter gets a non-instrumental benefit from voting for his favorite candidate. This intrinsic benefit is a random variable, equal to a constant $\bar{D}$ multiplied by a uniformly distributed variable between 0 and 1. Denoting the sincere voting benefit of voter i as D_i, we therefore have $D_i \sim U(0, \bar{D})$. The sincere voting benefit of each voter is private information but its distribution is commonly known to all voters.

Different from the pivotal voter model, there is a continuum of voters with measure one. The number of voters is infinite rather than finite for two reasons: First, the ethical voter models that I use as benchmark have a continuum of voters. Second, the continuum of voters assumption makes the model analytically tractable. I present a version of this model with a finite number of voters in Appendix A.7, in order to generate results on the effects of electorate size on third party vote shares.

[23]Making at least one of the turnout rates deterministic allows one to obtain a neat analytical solution without loss of intuition.

Since the number of voters is infinite, I define μ_j as the proportion of group j as a fraction of all voters for $j = 1, 2, 3$. As in the pivotal voter model, relative sizes of each group are public information. The parameter x, which represents the power difference between major parties, is now defined as the difference between μ_1 and μ_2, so that $\mu_1 = \mu_2 + x$.

As in the pivotal voter model, I restrict the analysis to the case with two strong parties and one weak party. Accordingly, I assume that μ_3 is much smaller than μ_1 and μ_2. The specific restriction is $\mu_3 < min(\frac{2b}{b+1}\mu_1, \frac{2b}{b+1}\mu_2)$, which ensures that party 3 never wins the election.[24] Therefore, voters in group 1 and group 2 always vote for their favorite parties, eliminating the possibility of a party 3 win.

Given these assumptions, the only decision problem is strategic vs. sincere voting for voters in group 3. The decision problem of voters in group 3 is to choose a cutoff point σ, where voters with $D_i < \bar{D}\sigma$ vote strategically (for party 2) and voters with $D_i \geq \bar{D}\sigma$ vote sincerely (for party 3). Hence, group 3's group-utility maximizing problem is as follows:

$$\max_{\sigma} \mu_3[yP(\sigma) + \int_{\sigma}^{1} \bar{D}z \, \mathrm{d}z] = \mu_3[yP(\sigma) + \bar{D}(\frac{1-\sigma^2}{2})]$$

[24]If voters in group 2 prefer party 3 over party 1 ($y > z/2$), an equilibrium can arise in which part of voters in group 2 and voters in group 3 coordinate on party 3 instead of party 2 (even when μ_3 is much lower than μ_2). That equilibrium can be eliminated by assuming that voters in group 2 prefer party 1 over party 3 ($y < z/2$). Even when that equilibrium exists, I will focus on the more plausible equilibrium where voters in group 2 and part of voters in group 3 coordinate on the majority party (party 2).

where $P(\sigma) = Pr(2 \text{ wins the election}|\sigma)$. The above objective function is the total expected utility of group 3 given the cutoff σ. The first term is the total expected instrumental gains of the group, given by the probability of the more favorable outcome (party 2 win) multiplied by the differential utility gain of having party 2 win over party 1.[25] The second term is the total non-instrumental benefits of the group. Since these benefits are received only by voters with $D_i > \bar{D}\sigma$, the lower bound of the integral is σ. Both of these terms are weighted by μ_3, the size of group 3.

The trade-off that voters in group 3 face when determining the cutoff σ is apparent from the objective function: A higher σ implies a higher chance that the more favorable party (party 2) wins the election, which increases the expected instrumental benefits of the group, but it also reduces the total non-instrumental benefits of sincere voting received by group members.

I explain the procedure for determining $P(\sigma)$ and describe the intermediate steps of the solution in Appendix A.5. The optimal cutoff level σ^* is found as follows:

$$\sigma^* = \begin{cases} 0 & \text{if } |x| > \frac{(1-b)(1-\mu_3)}{1+b} \\ min(\frac{(1+b)\mu_3}{(1-b)\frac{\bar{D}}{y}(1+x-\mu_3)}, 1) & \text{if } \frac{y}{\bar{D}} < \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2} \text{ and } |x| \leq \frac{(1-b)(1-\mu_3)}{1+b} \\ min(\frac{(1-b^2)\mu_3(1-\mu_3)+(1+b)^2\mu_3 x}{(1-b)^2\frac{\bar{D}}{y}(1+x-\mu_3)(1-x-\mu_3)+(1+b)^2\mu_3^2}, 1) & \text{if } \frac{y}{\bar{D}} \geq \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2} \text{ and } |x| \leq \frac{(1-b)(1-\mu_3)}{1+b} \end{cases}$$

[25]Remember that the probability of a party 3 win is zero, and hence the best possible outcome for voters in group 3 is a party 2 win. The utility of having party 1 win the election is normalized to zero.

Looking at the optimal cutoff, we first observe that strategic voting completely disappears ($\sigma^* = 0$) when the absolute value of x is large enough. The intuition is that as the election gets sufficiently one-sided, strategic voting by minority party supporters is unlikely to change the election outcome, so that minority party supporters refrain from costly strategic voting. We also have the following propositions, proved in Appendix A.6.

Proposition 2. *σ^* is (weakly) increasing in y and (weakly) decreasing in $\bar{D}$.*

This is the same qualitative result as the pivotal voter model, saying that the degree of strategic desertion of third parties increases with the stakes of the election (measured by y) relative to the benefit of sincere voting (measured by $\bar{D}$).

To visualize propositions 2 through 4, Figs.1.5,1.6,1.7 show the results of a numerical simulation exercise to graphically demonstrate how the extent of strategic voting (σ^*) and expected third party vote share (vote share of party 3) are affected by y, x and μ_3 respectively, using $y = 8, \bar{D} = 1, \mu_3 = 0.05, x = 0.025$, and $b = 0.2$ as benchmark parameters.[26]

Fig. 1.5 shows that σ^* is weakly increasing and the third party vote share is weakly decreasing in y. All party 3 supporters vote for party 2 (the third party gets zero votes) for high enough y.

Proposition 3. *σ^* is (weakly) increasing in x for $x < \tilde{x}$ and (weakly) decreasing in x for $x > \tilde{x}$, where $\tilde{x}$ is some threshold.*

[26]To make a fair comparison with the pivotal voter model of the previous section, these benchmark parameters are set to generate the same group sizes, mean turnout and fraction of strategic voters as the benchmark parameters of the pivotal voter model.

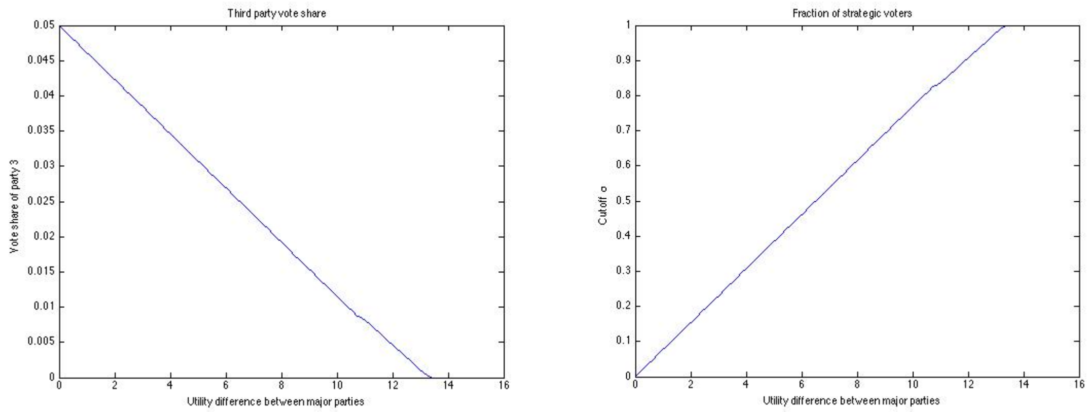

Figure 1.5: Third Party Voting vs. Utility Difference Between Major Parties - Rule-Utilitarian Model: (a) third party vote share, (b) fraction of strategic voters

Fig. 1.6 demonstrates that σ^* is increasing in x up to some threshold value of x ($\tilde{x}$ in Proposition 3) and decreasing afterwards. Conversely, the third party vote share decreases up to $\tilde{x}$ and then starts increasing. Thus, third party supporters strategically desert their favorite party more in closer elections. It should be noted again here that a slightly positive x indicates a closer election than $x = 0$ considering the strategic votes from party 3 supporters. This is the same qualitative result as in the pivotal voter model, that third party vote shares fall and the extent of strategic voting rises in closer elections.

Proposition 4. *σ^* is (weakly) increasing in μ_3 for $\mu_3 < \tilde{\mu_3}$ (some threshold value of μ_3), after which σ^* can increase or decrease with μ_3.*

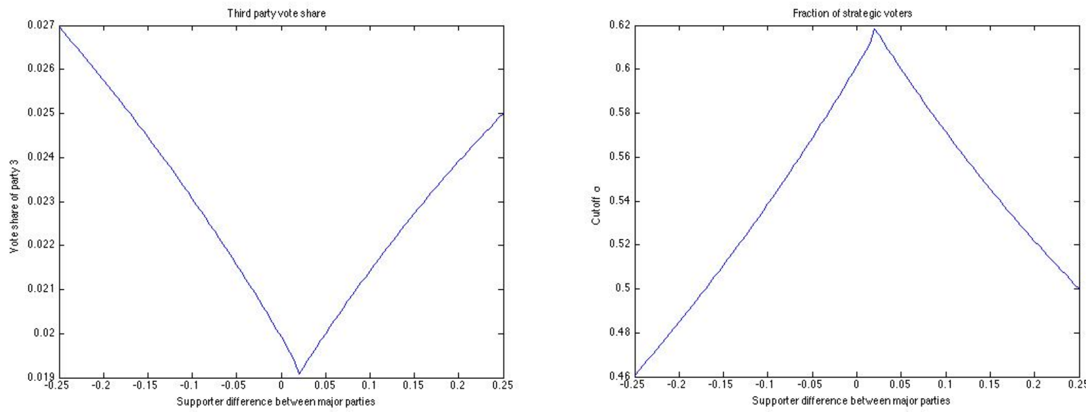

Figure 1.6: Third Party Voting vs. Difference in the Proportion of Supporters Between Major parties - Rule-Utilitarian Model: (a) third party vote share, (b) fraction of strategic voters

Fig. 1.7 shows that σ^* is increasing in μ_3 until σ^* reaches 1, and stays at the upper bound of 1 as μ_3 increases further.[27] There is an inverse U-shaped relationship between the third party vote share and μ_3, since the positive effect coming from the higher proportion of third party supporters is eventually counteracted by the negative effect on the third party vote share of higher σ^*.

Corollary. *Propositions 2-4 no longer hold when $D_i = 0$ $\forall i$ (so that there is no costly strategic voting).*

This corollary suggests that none of the comparative statics results of this model hold in the standard group-rule utilitarian model without costly strategic

[27]One should note here that the weakly increasing relationship between σ^* and μ_3 observed in Fig. 1.7 does not necessarily apply to all sets of parameter values. Proposition 4 suggests that an inverse U-shaped relationship between σ^* and μ_3 cannot be ruled out.

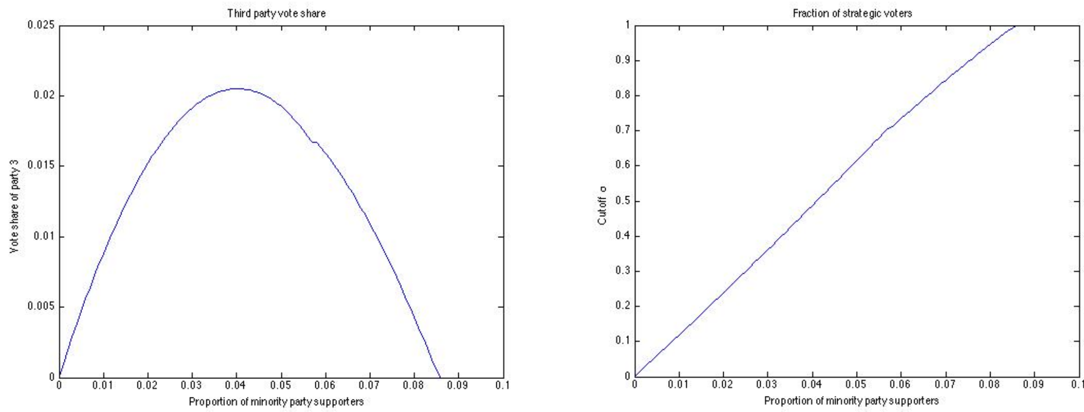

Figure 1.7: Third Party Voting vs. Proportion of Minor Party Supporters - Rule-Utilitarian Model: (a) third party vote share, (b) fraction of strategic voters

voting. To see why, suppose that non-instrumental benefits of sincere voting are not present so that $D_i = 0$ for all voters. Then the group-utility function of group 3 would only consist of instrumental benefits, given by the multiplication of y and the probability that party 2 wins. In that case, it is obvious that the optimal strategy of group 3 is that everyone votes for party 2 in order to maximize the probability of a party 2 win. That means that a weak third party with no chance of winning (party 3) would get zero votes and that the third party vote share would be independent of election-specific factors.

To summarize, the comparative statics implications of the rule-utilitarian model for strategic voting are roughly the same as the pivotal voter model with costly strategic voting. In both models, the extent of strategic voting is increasing and the third party vote share is decreasing in the utility difference between major parties, while the extent of strategic voting is higher and the third party vote share is lower in closer elections. The implications of the pivotal voter model on minority group size are not exactly the same but do not necessarily contradict the results of the rule-utilitarian model.[28]

The similarity of the main results between the pivotal and ethical voter models suggests that costly strategic voting rather than the type of voting model is driving the results. The addition of costly strategic voting generates predictions consistent with empirical evidence (positive third party vote shares and higher extent of strategic voting in closer elections) that baseline pivotal and ethical voter models without costly strategic voting are unable to explain.

Although the qualitative results of the pivotal and ethical voter models are very similar, the comparison of the models is imperfect since the pivotal voter model has a finite number of voters whereas the ethical voter model has a continuum of voters. If we were to assign both models a continuum of voters, then the pivotal voter model would yield no strategic voting at all since pivot events have zero probability when the number of voters is infinity and there are non-instrumental costs of strategic voting. All comparative statics implications would then change since the extent

[28] Even though Fig. 1.7 demonstrates a weakly positive relationship between σ^* and μ_3, Proposition 4 makes an inverse U-shaped relationship between σ^* and μ_3 (a result of the pivotal voter model provided in Appendix A.4) possible.

of strategic voting would always be zero and therefore would be independent of election-specific characteristics.[29]

One can make a case for either the pivotal or ethical voter models with costly strategic voting in terms of their ability to explain strategic voting in large elections: It could be the case that voters behave according to the group-utilitarian logic and engage in costly strategic voting since they follow the group rule, even though it is not individually rational for them to pay the non-instrumental cost of strategic voting. On the other hand, it could also be the case that voters behave according to the pivotal voter logic, but they engage in costly strategic voting even in large electorates where the probability of a pivot event is very small because they take the stakes of the election to be extremely high compared to their intrinsic cost of strategic voting, i.e. because $y/\bar{D}$ is extremely high.[30] It is difficult to judge which approach is more plausible in terms of explaining voter behavior without knowledge of $y/\bar{D}$ (which consists of two preference parameters that are difficult to estimate).

Although the qualitative results of pivotal and ethical voter models are mostly the same, a finite voter version of the ethical voter model predicts that third party vote shares can increase or decrease with the electorate size, depending on the relative strength of parties, while the pivotal voter model unambiguously predicts that electorate size is positively related to third party vote shares. Thus, the empirical

[29] This would essentially make the pivotal voter model equivalent to the sincere voting model, which (as I discussed in the introduction) fails to explain important features of observed voting patterns.

[30] As discussed in Section 1.2.6, other-regarding social preferences is one possible justification of an extremely high $y/\bar{D}$.

relationship between electorate size and third party vote shares can potentially distinguish between the two models. I illustrate the finite voter version of the ethical voter model and its comparative static predictions for electorate size in Appendix A.7.

1.4 Empirical Evidence

1.4.1 Existing Evidence

The first empirical fact to match is that third parties attract positive vote shares in simple plurality elections. The first two places in U.S. presidential elections have been held by Republican and Democratic party candidates continuously since 1920. However, third party candidates received an average of 5% of votes over that time period. The 2015 UK general election also provides interesting examples of substantial small party vote shares: For example, the UK Independence Party (UKIP) had a 12.6% vote share nationally despite winning only 1 of 650 seats available, by collecting substantial vote shares in many electoral districts where it is not in contention.

The second empirical fact to match is that the electorate includes both sincere voters (those who simply vote for the candidate they like most) and strategic voters (those who maximize expected instrumental benefits given the probabilities that each candidate wins). Recent papers by Kawai and Watanabe (2013) and Spenkuch (2013) estimate the ratio of strategic voters and they both find that sincere voting and strategic voting are both widespread across voters.[31] Therefore, data suggests that third party supporters will split between sincere and strategic voting, i.e. there is partial (as opposed to full or zero) strategic desertion of weak parties by their supporters.

Kawai and Watanabe (2013) and Spenkuch (2013) also find empirically that the extent of strategic voting is higher in closer elections. Since third party vote shares are inversely related to the extent of strategic voting, these results also imply that third party vote shares are lower in closer elections. Spenkuch (2013) also finds a higher extent of strategic voting when the election is perceived to be critical, which can be interpreted as evidence that the extent of strategic voting is higher when the stakes of the election are higher.[32]

Taking these as established empirical facts to match, I provide in Table 1.1 a summary of theoretical results from the models considered in this paper. As seen in Table 1.1, the addition of the mechanism implied by this paper (non-instrumental benefits to sincere voting, i.e. costly strategic voting) to the standard models generates theoretical predictions on third party vote shares that match the empirical

[31] Although their estimates on the extent of strategic voting is significantly different (Kawai and Watanabe (2013) estimates that around two thirds of voters are strategic whereas Spenkuch (2013) estimates that around one third of voters are strategic), the ratio of sincere and strategic voters are both far away from zero for both papers.

[32] Spenkuch (2013) finds a greater extent of strategic voting for the 2005 German federal election (compared to the 2009 German federal election). He describes the 2005 election as widely perceived to be a critical election since it followed a failed motion of confidence that triggered the dissolution of the Bundestag.

evidence whereas standard models without costly strategic voting cannot match these empirical findings.

Table 1.1: Summary Results of Different Models

	Strategic desertion of third parties	Impact of closeness on third party votes	Impact of stakes on third party votes	Impact of electorate size on third party votes
PVM, Standard	Full	None	None	None
PVM, Costly Str.	Partial*	Negative*	Negative*	Positive
RUM, Standard	Full	None	None	None
RUM, Costly Str.	Partial*	Negative*	Negative*	?
SVM	None	None	None	None

PVM: pivotal voter model, RUM: rule-utilitarian model, SVM: sincere voting model
Standard: costless strategic voting, Costly Str.: costly strategic voting
* Matches empirical findings

On top of the above facts, there are papers from the political science literature that also support the predictions of this paper on the issue of turnout. Using data from 1968 U.S. presidential elections, Bensel and Sanders (1979) finds that the highest percentage of non-voting is found on those who favor the minor party in their states, consistent with the prediction of this paper that minor party supporters are more likely to abstain. This point is worth mentioning particularly because there is no reason to expect such a result from a framework without costly strategic voting. The findings of Lacy and Burden (1999) that Ross Perot's candidacy increased overall turnout by three percentage points in 1992 U.S. presidential elections is also consistent with the effects of a popular third party candidate predicted by this paper.

1.4.2 New Evidence from U.S. Presidential Elections

The pivotal voter and rule-utilitarian models presented in this paper both yield the implication that third party vote shares are lower in closer elections. On the other hand, the pivotal voter model predicts that third party vote shares are higher when the electorate size is larger whereas the rule-utilitarian model generates ambiguous effects of electorate size on third party vote shares. I use data from the U.S. presidential elections from 1920 to 2012 to empirically test these implications.

The data source is Congressional Quarterly Voting and Elections Database and the dataset consists of over 75000 observations at the county level from 24 election years and 51 states. Given that the presidential race was between the Republican Party (R) and the Democratic Party (D) for all of these elections, I subtract R votes and D votes from the total number of votes, then divide this difference by the total number of votes to generate the sum of all third party vote shares for each county, which is the dependent variable.

Since winners are determined at the state level under the Electoral College system, the independent variables are closeness and electorate size at the state level. Closeness is the ratio of the losing major party (R or D) votes to the winning major party (R or D) votes at the state level. Therefore, closeness can take on values between 0 and 1, where higher values indicate closer elections. I exclude the observations in which the state-wide sum of third party votes is greater than either state-wide R or state-wide D votes. I also exclude the observations for which the state-wide third party vote share is zero. Electorate size is measured by state population.

Electoral vote closeness variable is the ratio of the number of Electoral College seats won by the losing major party (R or D) to the number of Electoral College seats won by the winning major party (R or D) at an election year. This variable is a measure of the national-level competitiveness of the election between R and D candidates. Incumbency is a dummy variable indicating whether an incumbent president is running. Regressions include county and year fixed effects along with state clustered standard errors.

Table 1.2 shows the results of the regression of third party vote shares on state closeness, state population and interaction terms. Closeness of the election between Republican and Democratic candidates has a negative effect on third party vote shares that is statistically significant at 1 percent, consistent with the theoretical predictions of the models. In terms of magnitude, a unit increase in closeness (going from the least close election to the closest election possible) causes third party vote shares to fall by around 2.5 percentage points. This is a sizable effect given that the average across the sample of third party vote shares is 4.8 percent.

On the other hand, state population has a negative but statistically insignificant impact on third party vote shares. This result does not support the prediction of the pivotal voter model that third party vote shares are increasing in electorate size, though it is not inconsistent with the predictions of the rule-utilitarian model. Interaction terms between state closeness, state population and incumbency are all statistically insignificant.

Table 1.2: Determinants of Third Party Voting, Panel Regressions

Dependent variable:
Third party vote shares

	(1)	(2)	(3)	(4)
State closeness	-2.373***	-2.519***	-3.098***	-2.782***
	(0.519)	(0.548)	(0.557)	(0.596)
State population	-0.109	-0.105	-0.118	-0.130
	(0.065)	(0.070)	(0.079)	(0.084)
State closeness*State population		-0.063	0.021	-0.012
		(0.151)	(0.142)	(0.149)
State closeness*Incumbency		2.455	1.945	1.441
		(1.548)	(1.663)	(1.406)
State population*Incumbency		-0.038	-0.037	-0.026
		(0.035)	(0.033)	(0.029)
State closeness*State closeness			-4.384**	
			(2.068)	
State closeness*				-4.771**
Electoral Vote Closeness				(1.970)
State population*				0.066
Electoral Vote Closeness				(0.04)
Constant	6.198***	6.080***	6.792***	6.400***
	(0.962)	(0.903)	(1.043)	(0.944)
R-squared	0.67	0.67	0.68	0.68
Observations	67,475	67,475	67,475	67,475

All panel regressions include year and county fixed effects. Standard errors are clustered at state. Robust standard errors are reported in parentheses. ***, **, * denote 1, 5 and 10% confidence levels, respectively. State population is measured in millions.

Moreover, the effect of state-level closeness on third party vote shares depends on whether the election is nationally competitive. Interaction term between state-level closeness and electoral vote closeness is negative and statistically significant at 5 percent, which suggests that closeness of the election at the state level matters much more when the election is close at the national level. Interpreting the coefficient, it suggests that state-level closeness has a modest impact of around 1 percentage point on third party vote shares for elections like the one in 1984 (where the election was extremely one-sided at the national level) and has an impact of around 6 percentage points (more than twice the average effect of 2.5 percentage points) in elections like the one in 2000 (where the election was extremely competitive at the national level).

The effect of state-level closeness on third party vote shares also depends on the level of state-level closeness, as suggested by the coefficient for the interaction term of state-level closeness by itself, which is negative and statistically significant at 5 percent. This means that the marginal impact of closeness on third party vote shares is higher in closer elections, i.e. that third party vote shares are decreasing and concave in state-level closeness. This is in line with what the pivotal voter model predicts on the shape of the relationship between third party voting and election closeness. Going back to the predictions of the pivotal voter model depicted in Fig. 1.2, one observes that increases in closeness create negligible effects on third party voting when the election is one-sided whereas increases in closeness create substantial declines in third party voting when the election is close.

1.4.3 Contiguous County Estimation

Panel data fixed-effects estimation is not particularly well-identified. Take the case of a local popularity shock for a third party candidate, which increases third party vote shares in that locality compared to other parts of the country. As the third party increases votes, he can both pull votes from the trailing major party (thus making the election less close) or the leading major party (thus making the election closer). Thus, local shocks to the popularity of a third party candidate can change both third party vote shares and closeness of the election in the locality for reasons unrelated to the strategic switching motive.

To address this concern, I employ the contiguous county estimation method used by Dube, Lester and Reich (2010). The idea is to compare neighbor counties in different states in terms of third party vote shares in order to control for differences in residual third party vote shares. This method provides better identification than the conventional fixed effects estimation to the extent that contiguous counties across a state border are better controls to each other than a randomly chosen county, since political preferences that affect the popularity of third parties are likely to be more similar across neighbor counties.

To implement this method, I first limit the sample to the list of contiguous pairs of counties provided by Dube et al (2010). I then estimate two specifications, one with county and time fixed effects and the other with country and pair-specific time fixed effects. For both specifications, standard errors are two-way clustered at state and border segment, where a border segment is defined as the set of all counties on both sides of a border between two states.[33]

Table 1.3 reports the results of the specification with county and time fixed effects. Since the earlier fixed effects estimation also used county and year fixed effects, this is essentially to check whether there is a significant change in regression coefficients due to the change in the sample. Comparing Table 1.3 to Table 1.2,

[33]For a detailed discussion of the methodology, see Dube et al (2010).

we see that the level and statistical significance of coefficients are very similar, indicating that the sample change does not significantly alter the results. State-level closeness is statistically significant at 1 percent and the interaction term between state-level closeness and electoral vote closeness is statistically significant at 5 percent for both samples. The only difference is that the coefficient for the interaction term of state level closeness with itself is statistically insignificant for the contiguous county sample.

Table 1.4 demonstrates the results of the specification with county and pair-specific time fixed effects. State-level closeness has a negative impact on third party vote shares that is statistically significant at 10%. The size of the impact is smaller compared to the specification with time fixed effects, with a unit increase in closeness inducing a decrease in third party vote shares of around 1.4 percentage points (compared to around 2.5 percentage points for the time fixed effects specification). Moreover, the interaction terms of state-level closeness with itself and with electoral vote closeness are statistically insignificant for the pair-specific time fixed effects specification.

Although the contiguous county estimation seems to produce weaker results than the regular panel data fixed effects estimates (lower impact of closeness on third party vote shares, statistically insignificant coefficients for interaction terms), one should interpret the results of the contiguous county estimation method with the caveat that it includes a very high number of fixed effect terms. Since fixed effects soak up a lot of the variation in the dependent variable (as confirmed by extremely high R-squares) and therefore sends regression coefficients towards zero,

Table 1.3: Contiguous County Estimation, County and Time Fixed Effects

Dependent variable:
Third party vote shares

	(1)	(2)	(3)	(4)
State closeness	-2.143***	-2.370***	-2.729***	-2.642***
	(0.540)	(0.598)	(0.672)	(0.663)
State population	-0.078	-0.068	-0.075	-0.090
	(0.066)	(0.066)	(0.070)	(0.074)
State closeness*State population		-0.116	-0.048	-0.071
		(0.150)	(0.133)	(0.148)
State closeness*Incumbency		2.450	2.113	1.400
		(1.624)	(1.689)	(1.414)
State population*Incumbency		-0.051	-0.051	-0.038
		(0.038)	(0.037)	(0.033)
State closeness*State closeness			-3.324	
			(2.375)	
State closeness*				-5.225**
Electoral Vote Closeness				(2.496)
State population*				0.071*
Electoral Vote Closeness				(0.039)
Constant	0.020	0.023	0.019	0.031
	(0.066)	(0.067)	(0.067)	(0.067)
R-squared	0.59	0.60	0.60	0.60
Observations	49,707	49,707	49,707	49,707

All regressions include year and county fixed effects. Standard errors are clustered at state and border segment. Robust standard errors are reported in parentheses. ***, **, * denote 1, 5 and 10% confidence levels, respectively. State population is measured in millions.

the true effect of closeness on third party vote shares, as well as the effects of the interactions of state-level closeness with itself and with national-level closeness, are stronger in all likelihood.

Table 1.4: Contiguous County Estimation, County and Pair-Specific Time Fixed Effects

Dependent variable:
Third party vote shares

	(1)	(2)	(3)	(4)
State closeness	-1.385*	-1.411*	-1.357*	-1.409*
	(0.765)	(0.771)	(0.755)	(0.773)
State population	-0.017	-0.016	-0.013	-0.021
	(0.055)	(0.055)	(0.055)	(0.056)
State closeness*State population		-0.007	-0.021	-0.008
		(0.102)	(0.108)	(0.103)
State closeness*Incumbency		-0.212	-0.194	-0.253
		(1.509)	(1.493)	(1.509)
State population*Incumbency		-0.004	-0.004	-0.002
		(0.019)	(0.019)	(0.018)
State closeness*State closeness			1.083	
			(2.016)	
State closeness*Electoral Vote Closeness				-0.365
				(1.546)
State population*Electoral Vote Closeness				0.015
				(0.023)
Constant	0.013	0.012	0.014	0.013
	(0.050)	(0.050)	(0.049)	(0.050)
R-squared	0.93	0.93	0.93	0.93
Observations	49,707	49,707	49,707	49,707

All regressions include year and pair-specific county fixed effects. Standard errors are clustered at state and border segment. Robust standard errors are reported in parentheses. ***, **, * denote 1, 5 and 10% confidence levels, respectively. State population is measured in millions.

1.5 Conclusion

Pure strategic voting or pure sincere voting ideas do not reasonably capture voter behavior. If voters were purely strategic, we would expect zero votes for weak third parties. If voters were purely sincere, we would expect no strategic desertion of third parties and third party vote shares to be independent of election closeness between front-runners. I present a hybrid approach based on the notion of non-instrumental benefits to sincere voting, where voters are strategically motivated in that they take into account the winning chances of each candidate to maximize expected utility, but also sincerely motivated for non-instrumental reasons. I show that this approach generates predictions consistent with existing empirical evidence that standard models cannot match.

Along with generating theoretical predictions consistent with existing empirical evidence, this paper presents new empirical evidence from U.S. presidential elections, which supports the prediction that third party vote shares are lower in closer elections. Moreover, the marginal impact of closeness on third party voting is stronger in closer elections, as predicted by the pivotal voter model. Empirical evidence also suggests that voters take into account national-level closeness as well as state-level closeness when deciding whether to desert third parties or vote for them.

This paper introduces the idea of costly strategic voting and explains its implications for third party voter behavior using a framework where the number of candidates, policy choices of candidates and the share of voters supporting each candidate are exogenous. A possible direction of future research is to build a framework including the mechanism introduced by this paper (non-instrumental benefits to sincere voting) where the candidate choices such as entry and policy decisions are endogenous. That framework would shed light on how the notion of costly strategic voting affects the decisions of politicians on whether to enter the race and on which policies to support.

Chapter 2: Electoral Strategies: Persuasion, Mobilization, Centrism

2.1 Introduction

What strategies should a candidate use in order to attract votes? There is a long-standing debate, for example, about whether elections are won by targeting "swing" voters or by "mobilizing one's base". The first are groups with a high concentration of voters who could swing to either candidate so that small changes in a candidate's position may yield a significant change in her vote share. The second (a candidate's "core voters") are groups who are likely to vote for a candidate if members of the group come out to vote, but need to be mobilized in order to turn out.

In this paper we investigate winning electoral strategies in terms of characteristics of voting groups and argue that answers to the question of whom to target are not so simple. One may get the impression from at least some of the literature that whether to target swing or core voters is an "either-or" question depending on the relative sizes of the groups and the cost of mobilizing the latter group. We will argue that the question of which groups to target in terms of their underlying characteristics in order to win an election is far more complex. In fact, for given characteristics of the electorate there may be multiple winning strategies – targeting swing voters, targeting core voters, and targeting a mixture of the two. Moreover, there will be "discontinuities" in electoral positions that win elections, for example where slightly away from center platforms on one side and extremist positions on the other side win elections, but centrist positions do not.

This model may be useful not only for illumination issues of specific electoral strategies – targeting swing versus core voters, centrism versus extremism, etc. – but also as a vehicle for analysis of other issues concerning electoral strategies. A key aspect of the model is the role of abstention, a factor that formal models of electoral strategies generally ignore.

We quickly note some questions we will not address in terms of electoral strategies. One is the issue of what Cox (2009) calls "coordination", meaning reducing the number of electoral competitors to increase one's vote share. Though this is important (consider multi-candidate primary elections) we consider the case where there are only two candidates. Second, we consider a single election, rather than a sequence of elections in which a candidate's strategy in one race may affect her electoral options in a subsequent race. This would be the case, for example, if the policy positions enunciated by a candidate in one race might limit what positions she could credibly take subsequently, as in Meirowitz (2005). It could also reflect changes in voter preferences as a result of the positions candidates took in previous races. We also do not consider possible electoral advantages (or disadvantages) of ambiguity in this paper (though we will briefly touch on why ambiguity may or may not be useful). Finally, we do not discuss explicit "vote buying", which is also important but not our focus. One could use the more general term of "clientelism" as an electoral strategy, but since in our view this term has many interpretations, we prefer not to enter into the question of whether our results bear on clientelism or not.

The paper is organized as follows. In section 2.2 we discuss related literature. Section 2.3 presents the basic model of attracting voters, which serves as a "workhorse" model for studying the issues raised here as well as others. Section 2.4 examines out some basic cases – centrism, mobilizing one's base, targeting swing voters, and presents baseline results. Section 2.5 considers the implications of voting groups differing over several dimensions, with a key result being that there may be more than one strategy that is consistent with winning the election for given characteristics of voting groups. In section 2.6 we consider the implications for of a candidate facing a very popular challenger. In this section we show conditions for the optimality of targeting a minority by choice of policy position. Section 2.7 adds extreme partisans to the model and studies the implications for voting strategies. In section 2.8 we further investigate how the interaction of different voting group characteristics can lead to multiple winning strategies for a candidate and how these may collapse depending on the type of opponent a candidate faces. Section 2.9 briefly discusses the implications of candidate uncertainty about the characteristics of voting groups. Section 2.10 summarizes the main results and concludes.

2.2 Literature

The academic literature on successful electoral strategies is obviously very large (and of course, the non-academic literature, if such a term is appropriate, is immense). We limit ourselves to the literature most directly related to the questions given above, as well as to some papers related to our general methodology.

2.2.1 Who Gets Targeted?

As Cox (2009) asks, "How do political parties allocate targetable goods – such as private goods targeted to individuals, local public goods targeted to geographic areas, or tax breaks targeted to specific industries or firms – in order to optimize their electoral prospects?" The most discussed issue is whether to target "swing" or "core" voters. Cox highlights this question and argues that there are two main 'camps' on this question – those who favor the "core voter model" of Cox and McCubbins (1986) and others versus those who favor the "swing voter model", for example in Lindbeck and Weibull (1987). Swing voters are generally defined as those who are close to indifference between candidates so that small changes in candidate's positions or other factors may easily swing them from one candidate to the other. It is often added that they are likely to vote, so that the issue is inducing them to vote in for a specific candidate rather than whether to vote. A swing group is one having a lot of voters in this situation, so that small changes in positions can induce large changes in vote shares from the group. Formally, there is a high density of such voters in the group. Cox reviews the literature and goes on to cite a significant number of papers that present evidence taken as supporting both views.

There are different definitions of core voters. One approach is that core voters are those who are "predisposed" to vote in favor of a party or candidate. This predisposition may be, for example, on programmatic grounds (Stokes (2005)) or on the basis of strongly held partisanship. A key issue is whether these voters are both strongly predisposed and very likely to vote (so that they can be safely ignored in choosing strategies to gain votes) or whether they must be mobilized to turn out for their favored candidate or party.

Dixit and Londregan (1996) present an alternative, though related, concept of a party's core voters, namely those voters with whom the party has an "advantage over its competition at swaying voters in a group with offers of particularistic benefits". Hence, it is not so much a 'predisposition' to vote for the party, but the party's ability to induce them to vote for it on grounds other than the party's policy. Under both definitions the party does not get their votes by changing its policy positions. Moreover, core voters are seen as those who will vote heavily for the party at presumably relatively low cost to the party per vote.

2.2.2 Modeling Voter Choices and Voting Equilibrium

Formal models of competition between two candidates are numerous. Even among those, there are several that focus on the issue of how group characteristics affect politician strategies, such as Lindbeck and Weibull (1987) or Dixit and Londregan (1996) already mentioned. One aspect is the importance for voting decisions of those characteristics that a candidate cannot change (or, as in the citizen-candidate model, policies he is believed certain to enact, campaign promises notwithstanding) versus things like those issues where a candidate can make a credible commitment to a position (or range of positions) that she chooses. The first may be termed "immutable" characteristics (or positions), the second "mutable" characteristics (or positions). For example Krasa and Polborn (2012) consider how the connection between the two affect equilibrium positions in a two-candidate race. Matakos and Xefteris (2017), considering policies towards income redistribution as in Roemer (1998), show how candidate equilibrium positions are affected by group characteristics such as size, marginal utility of income, and concentration of non-economic characteristics (which determines how 'swing' a group is).

A common result in these papers is that in the Nash equilibrium (if it exists) the two candidates converge to the same policy in equilibrium when for example, they are purely office motivated or if they are policy-motivated but have no uncertainty about the position of the median voter. Platform differences can emerge in equilibrium under uncertainty of this type, where they are driven by ideological differences of the candidates. These papers also assume no voting costs and hence no abstention – which would be generated by voters seeing no significant difference between candidate platforms – even though equilibrium in the models is often characterized by identical platforms of the two candidates.

The approach in this paper is quite different, in no small part because we feel that using the above type of model to study which groups are targeted by a candidate misses a number of issues that we think are central to this question. Are differences in candidates' platforms in equilibrium generated simply by their underlying ideology and uncertainty about voter preferences, or do they also reflect strategic choices? If a candidate has the ability to target groups more effectively than her opponent, how does she use that ability to gain votes? How does she successfully differentiate herself from her opponent to win an election? How important is mobilizing one's base to turn out to vote, a question that cannot be answered in models that assume full participation?

2.3 A Simple Model of Attracting Voters

We now set out a simple model of how a candidate's observed platform, summarized by a variable $\omega \in (0, 1)$, may attract voters. There are two candidates, denoted I and C, but we focus on candidate I who can better "signal" the policies she will adopt after the election, and we consider how she can use that ability to attract voters. This could be the incumbent who has the ability to use government expenditures to target voters. In the literature (for example, Cox and McCubbins (1986), Lindbeck and Weibull (1987)) the question of which voters are targeted in order to win elections has been modeled in terms of distribution of a public good subject to a budget constraint, where candidates commit to a post-electoral distribution. Alternatively, the composition of government expenditures may be used to signal the incumbent's priorities even if commitment is not assumed, as in Drazen and Eslava (2010)[1].

Since we assume that the two main groups prefer either $\omega = 0$ and $\omega = 1$, and that platforms are binding commitments (though may only give a range of possible ω), any choice of ω by a candidate is equivalent to a decision on distribution of a

[1]There is a good bit of evidence that the composition of spending changes before elections to attract votes both at the local and national levels. See Brender and Drazen (2013) for a summary and evidence on the national level across countries.

government-provided good.[2] ω may also be thought of as the tax rate, consistent with the literature on electoral determination of tax policy, where the rich prefer a tax rate on income of 0 and the poor prefer (absent disincentive effects) a tax rate of 1.

We will sometimes refer to candidate I as the "incumbent" and candidate C as the "challenger", but of course it may be the challenger who can more precisely or credibly indicate her platform ω (though the interpretations of ω in the previous paragraph may be consistent with it being the incumbent who can do this). The main point is that our analysis concentrates on the candidate who can make more precise policy statements and, given this ability, the strategies she may use to win elections, and we call this candidate I and her opponent C for ease of exposition.

2.3.1 Voter Utility

For the bulk of the paper we assume there are two types or groups of voters[3]: group A, who favor ω as high as possible, group B, who favor ω as low as possible. Voters also have candidate-specific or "partisan" preferences which are independent of ω or any actions the candidate may take. We denote voter i's "partisan" preference for candidate $P \in \{I.C\}$ by π_P^i. The utility of individual i in group $h = A, B$ if the candidate P is elected and implements policy ω may be represented as

[2] Cox (2009) labels such commitment "outcome-contingent transfers" that is "promising to deliver benefits if and only if one wins"

[3] In some of our analysis, we will introduce a third group O whose most preferred position is in the center, namely $\omega = \frac{1}{2}$.

$$
\begin{aligned}
u_A^i(\omega; P) &= \ln \omega + \pi_P^i \\
u_B(\omega; P) &= \ln(1-\omega) + \pi_P^i
\end{aligned}
$$

Since it is only net partisan preference that matters, we denote by $\lambda^i = \pi_C^i - \pi_I^i$ voter i's net partisan preference for candidate C, so that a voter with $\lambda^i < 0$ has a preference for I independent of ω, while one with $\lambda^i = 0$ has no partisan preferences. A voter with λ^i high enough in absolute value would never vote for one of the candidates no matter what the ω positions of the two candidates were.

We note that the formulation where candidates can change their positions on some issues ω but have immutable characteristics (represented by the π_P^i) is itself pretty standard. One may think of this partisan factor as representing a "citizen-candidate" (Osborne and Slivinski (1996), Besley and Coate (1997)) aspect of candidates' positions, that is those policy aspects that voters believe will be carried out by a candidate independent of any campaign promises he or she might make. Hence the model can capture fully credible platform commitments by a candidate, partially credible commitments (where the incumbent can commit only to a range of ω), and aspects where voters know (or at least believe they know) what a candidate will do if elected, so that campaign statements would have no effect.

A key group characteristic is the distribution of partisan preferences within a group as in Dixit and Londregan (1996). We consider two possible distributions of the λ^i in a group. One is that the distribution of candidate preference is normal with mean $\bar{\lambda}_h$ and standard deviation σ_h for $h = A, B$. Without loss of generality we will consider group A to be the incumbent's "natural" or core constituency and group B to be the challenger's core constituency. That is, we take $\bar{\lambda}_A < 0$ and $\bar{\lambda}_B > 0$ (though individual-specific λ_A^i and λ_B^i are normally distributed around these means). σ_h measures the concentration of partisan preferences in the group ("ideology" in Dixit and Londregan, "within-group homogeneity" in Matakos and Xefteris (2017)).

The other possibility is meant to represent two subgroups within each group, one centered around a $\bar{\lambda}_h$ that is low in absolute value ("moderate" partisans), the other around a $\bar{\lambda}_h$ that is high in absolute value ("extreme" partisans), with known proportions of extremists and moderates in each group. This will be investigated starting in section 2.7.1.

2.3.2 Voter Information

We can generally represent the information that voters have by a probability distribution of possible ω for both candidates, denoted by density functions $\psi^I(\omega)$ on I and $\psi^C(\omega)$ on C. We assume, for simplicity, that all voters have the same information set, though we do not restrict the distributions ex ante. Our key assumption that I has an advantage of greater ability to signal her position ω can be represented by the distribution $\psi^I(\omega)$ being "tighter" than $\psi^C(\omega)$. We consider the polar case in which I's ω is known while C's ω is uniform over one of three ranges $\mathbb{C}^L = \left[\omega^l, \frac{1}{2}\right]$, that is, a "leftist" challenger; $\mathbb{C}^R = \left[\frac{1}{2}, 1-\omega^l\right]$, that is, a "rightist" challenger; and $\mathbb{C}^M = \left[\omega^l, 1-\omega^l\right]$, that is a challenger on neither side of the spectrum, where the distribution of possible preferences is centered on $\frac{1}{2}$.[4] Note crucially that "left" and "right" refer to positions on the ω-line. This corresponds to real world notions of left and right if ω were a metric of a right-wing policy (such as fraction of the budget devoted to guns rather than butter), but we could have ω refer to a left-wing policy, so that a higher ω (that is, on the right part of the line) is a left-wing policy. This is useful for some of our examples, as we will see below.

2.3.3 Voting Costs

Modeling abstention based on voting costs faces the question of why individuals bother to vote in a large electorate, as their probability of being pivotal approaches zero. Addressing this question is difficult and beyond the scope of this paper. We take it for granted that some citizens never vote no matter what the positions of the candidates. We do not include them as part of the electorate (though conventional measurement counts them as voters who abstained) and consider only voters whose decisions are affected by candidates' positions and costs of voting. For these 'potential' voters, we simply assume that an individual abstains when his voting cost outweighs the difference in utility expected from the two candidates. That is, each voter views himself as pivotal in deciding whether or not to vote.[5] We assume that each voter has a cost of voting $\gamma^i \geq 0$. For much of the paper we assume that all individuals have the same voting cost $\gamma^i = \gamma > 0$, where the likelihood that

[4] To bound utility away from $-\infty$, we assume in the computations that ω^l is close to but strictly greater than 0.

[5] An alternative approach would be that the non-instrumental benefit of voting could be candidate specific, as in the first chapter of this dissertation.

a given voter will abstain can still vary over voters depending on their λ^i and of course the position ω of the candidates.

We will first solve for vote shares as a function of candidate policies, partisan voter preferences and costs of voting where we assume that there are no exogenous shocks that affect turnout. We will label this as intention to vote. We will then add a random turnout shock that will convert vote shares to winning probabilities but will still allow use of the same formal model to analyze electoral strategies.

A voter i in group A intends to vote for I only if the difference in expected utility under I and C is at least as large as the cost of voting:

$$E(\ln \omega|I) - E(\ln \omega|C) - \lambda_A^i \geq \gamma,$$

where $E(\ln \omega|I) \equiv \int_\omega \ln \omega \psi^I(\omega)\, d\omega$ and analogously for $E(\ln \omega|C)$ with $\psi^C(\omega)$ replacing $\psi^I(\omega)$. Similarly, he intends to vote for C if the expected utility gain from having C rather than I elected is at least as great as the cost of voting:

$$E(\ln \omega|C) - E(\ln \omega|I) + \lambda_A^i \geq \gamma$$

Finally, a member of group A plans to abstain rather than turn out to vote for one of the candidates when the difference in his utility under the two candidates is less than the voting cost γ^i so that (reversing the two inequalities above)[6]

$$-\gamma < E(\ln \omega|I) - E(\ln \omega|C) - \lambda_A^i < \gamma$$

[6] Remember that some citizens never vote and are excluded from the analysis as discussed above, while some voters plan to vote but an exogenous shock, such as weather, may induce them to stay home. This will be discussed below.

Analogous equations hold for members of group B but with $\ln(1-\omega)$ replacing $\ln\omega$ (with individual-specific voting costs, the γ would be replaced by γ^i, but one can immediately see why individual-specific λ^i can have similar effects).

2.3.4 Vote Shares and Election Outcomes

We may then write the fractions of voters in group A who intend to vote for I as

$$v_A^I = F_A\left(\int_\omega \ln\omega\psi^I(\omega)\,d\omega - \int_\omega \ln\omega\psi^C(\omega)\,d\omega - \gamma\right)$$

where $F_A(\cdot)$ is the CDF of a standard normal with mean $\bar{\lambda}_A$ and standard deviation σ_A. In the case where I's position in known, the first term in parentheses would simply be $\int_\omega \ln\omega\psi^I(\omega)\,d\omega = \ln\omega$. Analogously, the fraction who intend to vote for C is

$$v_A^C = 1 - F_A\left(\int_\omega \ln\omega\psi^I(\omega)\,d\omega - \int_\omega \ln\omega\psi^C(\omega)\,d\omega + \gamma\right)$$

Finally the fraction of group A who plan to abstain given the candidate's positions is $\varnothing_A = 1 - v_A^I - v_A^C$, which could be written

$$\begin{aligned}\varnothing_A = {} & F_A\left(\int_\omega \ln\omega\psi^I(\omega)\,d\omega - \int_\omega \ln\omega\psi^C(\omega)\,d\omega + \gamma\right) \\ & - F_A\left(\int_\omega \ln\omega\psi^I(\omega)\,d\omega - \int_\omega \ln\omega\psi^C(\omega)\,d\omega - \gamma\right)\end{aligned}$$

The vote shares in group B would be analogous, but with $\ln(1-\omega)$ replacing $\ln\omega$ and with $F_B(\cdot)$ replacing $F_A(\cdot)$, where $F_B(\cdot)$ is the CDF of a standard normal with mean $\bar{\lambda}_B$ and standard deviation σ_B.

If all voters who intend to vote actually do vote, candidate I's share of votes is

$$\begin{aligned} S &= \frac{\upsilon_A^I \varphi_A + \upsilon_B^I \varphi_B}{\upsilon_A^I \varphi_A + \upsilon_B^I \varphi_B + \upsilon_A^C \varphi_A + \upsilon_B^C \varphi_B} \\ &= \frac{V^I}{V^I + V^C} \end{aligned}$$

where φ_A and φ_B are the fractions of the two groups in the population and where $V^I = \left(\upsilon_A^I \varphi_A + \upsilon_B^I \varphi_B\right)$ and $V^C = \left(\upsilon_A^C \varphi_A + \upsilon_B^C \varphi_B\right)$. For simplicity, call this share S if intentions are fulfilled simply the "vote share".

To convert vote shares into probabilities of winning, suppose that a shock on voting day ("bad weather") implies that only a fraction τ^I of candidate I's voters actually fulfill their intentions to turn out and vote for her. Similarly only a fraction τ^C of candidate C's voters fulfill their intentions to turn out and vote for her. Suppose these fractions τ^I and τ^C are independent of any voter or candidate characteristics, as well as unknown ex ante by candidates or voters. Hence the ratio τ^I/τ^C is a random variable. The threshold for winning (getting 50% of the actual vote) would not be $S = \frac{1}{2}$ but

$$\frac{\tau^I V^I}{\tau^I V^I + \tau^C V^C} = \frac{1}{2}$$

which could be written $\tau V^I = V^C$where $\tau \equiv \tau^I/\tau^C$. This implies that at the threshold we can write

$$S = \frac{V^I}{V^I + \tau V^I} = \frac{1}{1+\tau}$$

where $\frac{1}{1+\tau}$ is a random variable, say with CDF $\Upsilon(\cdot)$ and mean of $\frac{1}{2}$. We may then write the probability of winning as a function of S, the "vote share", as

$$\Pr\left(S \geq \frac{1}{1+\tau}\right) = \Upsilon(S)$$

where the probability of winning is monotonically increasing in S. Hence, though vote share is stochastic we can analyze I's electoral strategies in terms of S defined as $\frac{V_I}{V_I+V_C}$, which I can take as non-stochastic if she knows demographic characteristics, relative to a stochastic winning threshold $\frac{1}{1+\tau}$, which is fully exogenous to the candidates. We shall argue that candidate I may have different ways of increasing S above a given level of to satisfy this condition for given characteristics of the electoral population: relative group size φ_A and φ_B; fraction of extremists in a group ϵ_A and ϵ_B; voter information on candidates' policies $\psi^I(\omega)$ and $\psi^C(\omega)$; average partisan preferences $\bar{\lambda}_A$ and $\bar{\lambda}_B$; dispersion of partisan preferences σ_A and σ_B; and voting costs γ.

2.3.5 Candidate Behavior

Our interest in the paper is to investigate winning strategies for a candidate I who can credibly commit to a policy position ω given the characteristics of the electorate and the type of challenger she faces. Because we want to focus on how voting group characteristics affect I's position relative to C, we assume that I has no ω preferences and cares only about being elected.

Hence, as already discussed in section 2.2.2, we focus on the choice of ω by candidate I for different ranges of choices by candidate C. One may interpret this as a situation where I faces a challenger who can neither change what voters believe is his general policy orientation nor make it as precise as I can. We will however note what the Nash equilibrium is when both candidates choose strategies simultaneously.

2.3.6 Defining "Swing" and "Core" Groups

As discussed in section 2.2, swing voters are often defined as those who are close to indifference between candidates so that small changes in candidate's positions or other factors may easily swing them from one candidate to the other. Although this would seem to imply a λ^i close to 0, a voter with a non-zero but not too large λ^i will be swing for some value of ω. A swing group is one with a lot of swing voters. In our model, this corresponds to a group with a high σ_h, that is where voter's partisan positions are very concentrated, so that there are a large fraction of voters in the group who will change their votes in the same ω range. One should add that in the presence of positive voting costs ($\gamma > 0$), a voter never moves directly from voting from one candidate to the other as the incumbent's ω changes, but always moves from a candidate to abstention and then to the other candidate with changes in ω. Only when $\gamma = 0$ (so the voter is certain to vote) does a voter swing directly from one candidate to the other for a marginal change in ω.

Core voters are those who are "predisposed" to vote in favor of a party or candidate, which in our model corresponds to a large absolute value of λ^i. A core group would then be one which satisfies either of two conditions. Either it is unimodal in partisan preferences and characterized by a large absolute value of $\bar{\lambda}_h$ – negative if the group is candidate I's base, positive if it is candidate C's base – so that the average voter in the group has a strong predisposition towards one candidate or the other. This would be combined with a sufficiently small σ_h, so that many voters in the group are characterized by a λ^i close to $\bar{\lambda}_h$. Alternatively, it can be a group with bimodal (or multimodal) distribution of the λ^i with a large fraction of "extremists", that is, those with high λ^i (in absolute value), highly concentrated around their average λ.

Note that relating the notions of swing and core voters to parameter values makes clear that a group could be more or less swing and more or less core. This possibility will figure in to the existence of multiple winning regions.

2.4 Some Basic Cases

We begin by illustrating some basic cases: the value of centrism when voting groups have opposing policy preferences but are otherwise identical; targeting swing voters when the two groups differ in how concentrated their preferences are; mobilizing one's base when there are voting costs and one's base is relatively unmotivated to vote.

2.4.1 Centrism (and the Effect of Concavity)

Politicians who espouse centrist policies will win elections against those who favor non-centrist policies if voters are concentrated around the center. Our locational argument does not however require such an assumption. When voters are symmetrically located away from the center, the same locational argument will hold in the absence of partisanship.

To demonstrate the role of concavity in voter preferences in inducing centrism, we start with the most basic case of the model with equal-sized groups ($\varphi_A = \varphi_B$), no average partisanship for either group ($\bar{\lambda}_A = \bar{\lambda}_B = 0$), and no voting costs ($\gamma = 0$). The dispersion of partisan preferences is assumed equal across groups, with $\sigma_A = \sigma_B = 0.2$. We further assume that I's ω is known while the C's ω is uniform over a range $\mathbb{C}^M = [.05, .95.]$.

Note first that if voters had linear preferences, they would compare the expected value of ω^C, which is 0.5, to I's ω in choosing how to vote. Given the symmetry of the two groups, any ω chosen by I would give her exactly 50% of the vote, but no more.

Things look different when voters have concave preferences over policy ω as we assume. The top panels of Fig. 2.1 plots I's vote share (panel a) and vote totals (panel b). We see that her vote share is maximized at the center, i.e. $\omega = 0.5$. This is due to concavity in voter preferences. As I moves away from the center to favor one of the groups, the utility gain of the favored group is less than the utility loss of the unfavored group so she loses the votes of the unfavored group faster than she gains votes from the favored group.

One may note that concavity of voter preferences implies that ambiguity on the part of I about her ω position cannot help the incumbent if voters have unbiased perceptions about I's actual position given her (ambiguous) policy announcement. If voters interpret ambiguous positions in a biased way, that is, by overweighting the possibility that the policies they will adopt if elected are those that they favor, we find in preliminary research that this may only help I if the ambiguity is fairly small.

One should note that what is important is centrism relative to the challenger. I's winning strategy would shift to the left if the challenger C were left-wing, that is, where C's possible positions are in $\mathbb{C}^L = \left[\omega^l, \frac{1}{2}\right]$ (Fig. 2.1, panel c), and to the right if C were right-wing with policies in $\mathbb{C}^R$ (Fig. 2.1, panel d). Note further that I's winning regions overlap when ω is close to $\frac{1}{2}$ in the three cases of a left-wing, center, and right-wing challenger. Hence, if I did not know the type of a challenger she would face when choosing her policy (that is, in a "simultaneous move game"), she could still guarantee victory by choosing a sufficiently centrist strategy.

Concavity of voter preferences also implies a "known-type" advantage for I, who gets 50 percent vote share or more for any ω, and she gets much more than 50 percent for ω close to 0.5. I can perfectly indicate her position, and thus can win easily over C if she chooses $\omega = 0.5$, the mean C's position, for whom only the distribution of possible ω is known. To the extent that it is incumbency in office that gives I the ability to more credibly or precisely communicate her post-electoral policy to the electorate, this is a type of incumbency advantage.

One should note however that incumbency may provide a disadvantage in terms of I choosing a position ω that maximizes her vote share or probability of winning. Holding office often requires making decisions that indicate specific position, and these decisions thus limit the positions that an incumbent subsequently running for re-election can credibly take. This can be represented by the possibility that I may be able to take positions only in some subregion of the ω line. We discuss this in section 2.8.3.3 below.

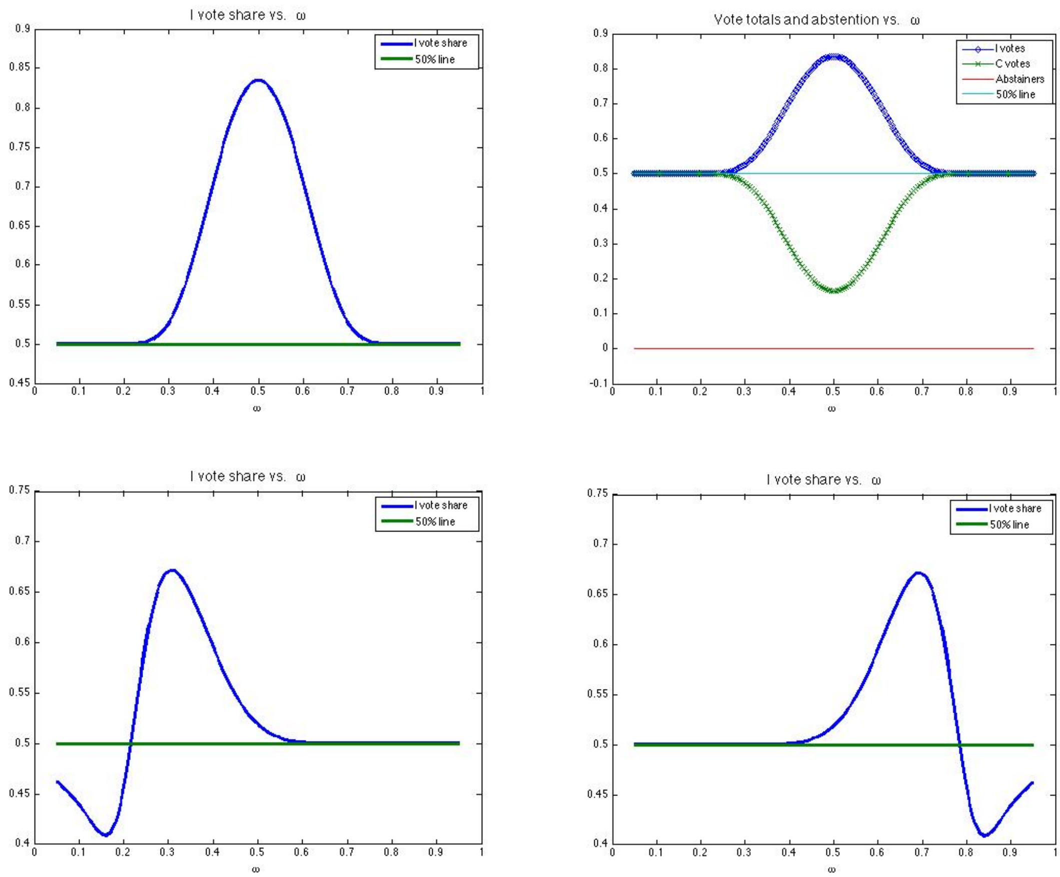

Figure 2.1: Centrism: (a) vote shares - centrist challenger, (b) vote totals - centrist challenger, (c) vote shares - leftist challenger, (d) vote shares - rightist challenger

2.4.2 Targeting Swing Voters (the Effect of Differential Preference Dispersion)

A standard view is that when turnout is not an issue swing groups will be targeted because doing so will deliver a large number of votes. We can represent this idea by taking the above case of equal-sized groups with no average partisanship for either group and no voting costs ($\gamma = 0$), but suppose that the dispersion of partisan preferences is different across groups, taking $\sigma_A = 0.1$ and $\sigma_B = 0.6$. In other words suppose that groups are identical in their demographic characteristics and have no average bias towards one candidate or the other, but that group A (those favoring high ω) is much more concentrated in their λ^i around 0. Suppose as before that I is facing a challenger with ω uniform over the range $\mathbb{C}^M = [.05, .95.]$ with a mean of 0.5.

In Fig. 2.2 (panel a) we see that I's vote share curve has shifted to the right, that is, towards higher ω. There is targeting of the more swing group. Panels b and c of Fig. 2.2, giving voting behavior of groups A and B respectively, make clear why this is so. In the absence of any partisan preferences ($\lambda^i = 0$ for all voters), concavity of preferences means that all group A voters would be indifferent between I's known policy and C's at the same point $\hat{\omega}^A = 0.416$ (defined by $\hat{\omega}^A = \int_\omega \ln \omega \psi^C(\omega)\, d\omega$), while all group B voters would be indifferent at $\hat{\omega}^B = 0.584$. Dispersion of the λ^i around $\bar{\lambda} = 0$ implies that there is some dispersion of indifference point around these respective values of $\hat{\omega}$, so that voters switch from one candidate to the other at different values of ω in the neighborhoods of the respective $\hat{\omega}$. The less dispersed (or

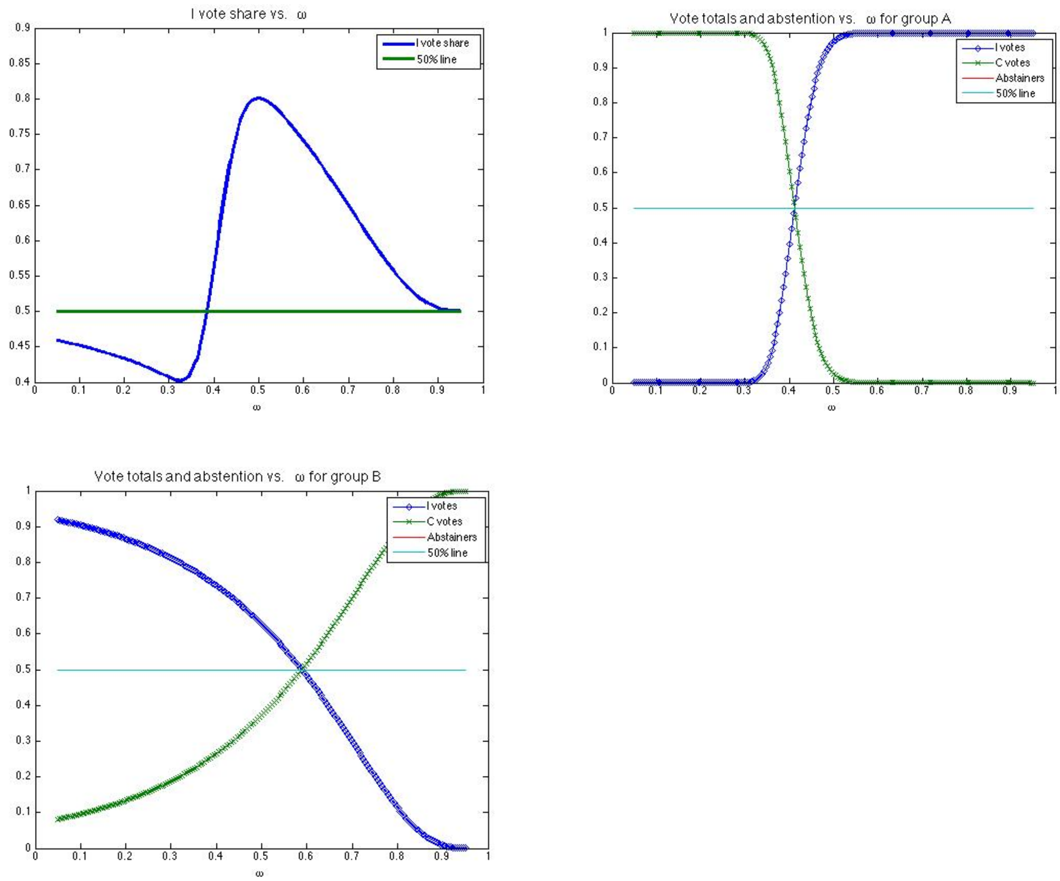

Figure 2.2: Targeting Swing Voters: (a) vote shares - centrist challenger, (b) vote totals vs. ω for group A, (c) vote totals vs. ω for group B

more concentrated) are voter preferences, the steeper are the curves of vote switching as illustrated by these two panels. When, as in our example, group A voters are more concentrated I gains votes from them much faster than she loses votes from the less concentrated group B voters, so her vote-maximizing policy shifts in their direction.

2.4.3 Mobilizing One's Base (the Effect of Preference Intensity and Voting Costs)

Conventional wisdom is that an alternative to winning elections by swinging likely voters is "mobilizing one's base" to come out and vote. Obviously this requires there to be positive voting costs ($\gamma > 0$) if some voters choose not to vote when comparing the expected utility difference under I versus C. As discussed above, we think of a candidate's base as those voters who, were they to come out and vote rather than abstain, are known to likely vote for that candidate. This could obviously be represented by an average λ in a group.

To illustrate the basic ideas, let's suppose the majority of the electorate has partisan preferences favoring I, but that the minority that favors C has more intense candidate preferences and thus are more motivated to vote. Hence, I needs to mobilize enough of her supporters to turn out to vote in order to offset a group of voters certain to prefer the opposing candidate. To represent this case, suppose that group A forms 55% of the electorate ($\varphi_A = .55$) but $\bar{\lambda}_A = -0.2$ while $\bar{\lambda}_B = 2$, meaning that group B voters have far stronger preferences towards C than group A has towards I. Suppose that both groups have the same level of concentration, say $\sigma_A = \sigma_B = 0.2$.

Fig. 2.3 illustrates the phenomenon in terms of vote totals for voting costs $\gamma = 0$ (panel a), $\gamma = 0.3$ (panel b),and $\gamma = 0.6$ (panel c) when C is neither right nor left wing, that is, her positions are distributed uniformly over the entire range $\mathbb{C}^M = [.05, .95.]$. One sees that as voting costs increase, I needs to move more towards the preferred policy of her base in order to increase her vote share above 50%. Panels b and c of Fig. 2.3 make clear that I is increasing her vote totals not by inducing group B voters to swing towards voting for her (Group B is so heavily disposed towards the challenger C in terms of λ preferences that they fully vote for C for any ω) but by shifting A voters from abstention to turning out.

The strategy of mobilizing one's base will not however work against a challenger who is on the same side of the ω spectrum as her (weakly-motivated) base. That is, consider the above case in terms of voting group characteristics, but suppose that I faces a right-wing challenger with $\omega \in \mathbb{C}^R = [.5, .95.]$ – that is, a challenger whose expected ω policy position coincides with the preferred policy of group A, I's base. As voting costs rise (hence making turnout of A voters more difficult) I must move farther and farther to the right. For high enough costs she will almost definitely lose, as illustrated in panel d of Fig. 2.3 where the challenger is in $\mathbb{C}^R$ and $\gamma = 0.6$. That is, even with the highest possible ω, I's vote share does not exceed 25%. The combination of the intensity of the challenger's candidate-specific support (from group B voters) and her "right-wing" policy stance attracting group A voters dooms the incumbent. This represents the problem of running against a candidate on the same side of the ω policy spectrum who has intense candidate-specific support.

A policy of mobilizing one's base also appears when the two groups differ in terms of the level of voting costs γ rather than in their $\bar{\lambda}$. To see this, now suppose that average partisanship is the same in absolute value across groups, say $\bar{\lambda}_A = -1$ and $\bar{\lambda}_B = 1$. Suppose that group B has no voting costs ($\gamma = 0$) but consider different

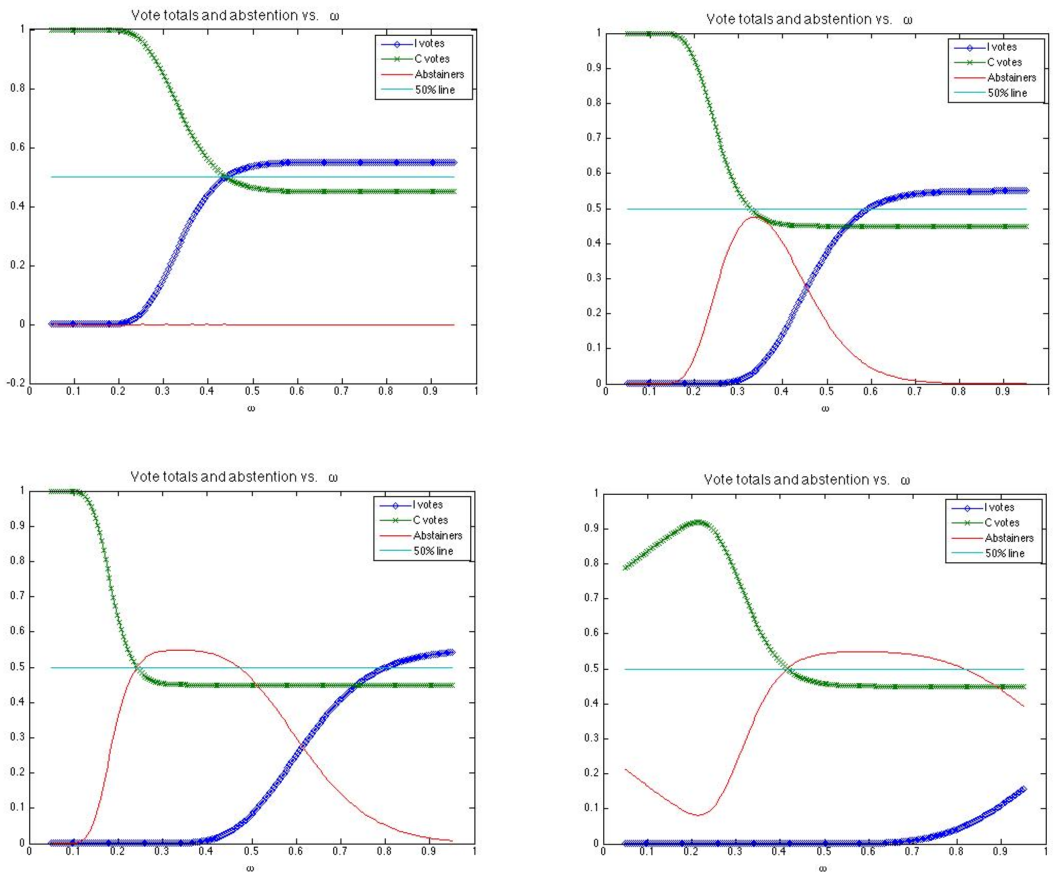

Figure 2.3: Mobilizing One's Base - Partisanship (a) vote totals - centrist challenger, $\gamma = 0$, (b) vote totals - centrist challenger, $\gamma = 0.3$, (c) vote totals - centrist challenger, $\gamma = 0.6$, (d) vote totals - rightist challenger, $\gamma = 0.6$

voting costs for group A. Fig. 2.4 plots I's vote share for $\gamma_A = 0$ (panel a), $\gamma_A = 0.75$ (panel b), and $\gamma_A = 1.5$ (panel c). We observe that heterogeneity across groups in terms of voting costs is another reason that induces a candidate to 'mobilize her base'. When there is no voting costs, I can win by being a centrist or even by favoring group B. As voting costs for group A rise, however, I needs to favor her base more and more to convince them to turn out to vote, that is, to favor her base in order to win the election.

2.5 Voting Groups Differ Across Multiple Dimensions

The base cases discussed in section 2.4 suggest that there is an intuitive translation from characteristics of voting groups to electoral strategies. When voting groups are similar except for their preferred policies and are likely to turn out to vote, then centrism is a winning strategy. When however, turnout is a problem among a candidate's base, centrism against a centrist challenger may simply induce significant abstention, so that a candidate may lose to a challenger with more motivated supporters. In such a situation, a candidate may need to move away from the center towards the preferred position of her base in order to motivate them to turn out so she can win. The base cases also gave support to a strategy of targeting voting groups that are very "swing" in that small changes in position can induce large shifts in their voting. Given group size, concentration of preferences within the group would be a key determinant of the success of such an electoral strategy.

If voting groups differed from one another only in a single dimension – average policy preferences, within-group concentration of policy preferences, motivation to

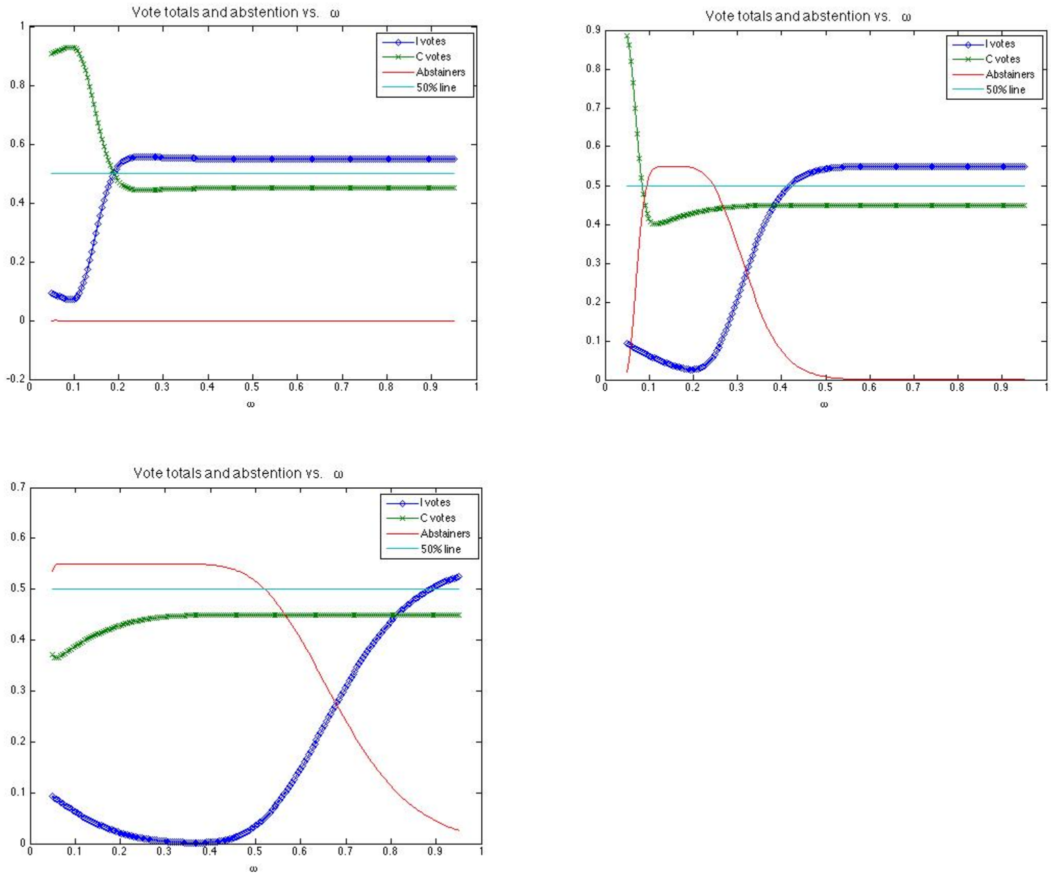

Figure 2.4: Mobilizing One's Base - Differential Voting Costs: (a) vote totals - centrist challenger, $\gamma_A = 0$, (b) vote totals - centrist challenger, $\gamma_A = 0.75$, (c) vote totals - centrist challenger, $\gamma_A = 1.5$

vote – then it makes sense that there is a single way to win elections depending on which dimension is most pronounced. For example, suppose voter preferences are similarly dispersed among those who favor right-wing versus left-wing policies and those groups are similar in size, but motivation to vote is low. The candidate who wins would be the one better able to mobilize her base, where differential motivation across core voters may be central.

However, suppose more realistically that groups differ in several dimensions, some of which might suggest targeting swing voters, others perhaps targeting one's core voters. Electoral strategies will thus depend on the interaction of differences in partisan preferences – both average intensity and concentration or dispersion – and positive voting costs. We now consider this in greater detail to show that the interaction of factors can lead to several phenomena which may seem counterintuitive ex ante but can be explained by a formal model that separates these factors.

2.5.1 "Dual" Electoral Strategies

A key result when we look at groups differing in multiple dimensions is that different electoral strategies may be consistent with winning elections. However, as we argue below, they cannot be simply associated with targeting swing voters or mobilizing one's base.

Suppose groups differ in both their average intensity of partisan preferences and in the dispersion of these preferences. Suppose that one group has stronger average preferences towards its preferred candidate, but that these preferences are more dispersed within the group. For example, suppose that $\bar{\lambda}_A = -0.1$ and $\bar{\lambda}_B =$

0.3 (group B has stronger average partisan preferences), but that $\sigma_A = 0.05$ and $\sigma_B = 0.3$ (partisan preferences are very concentrated in group A while they are significantly dispersed in group B, so that all members vote almost identically at any value of ω, i.e. as a bloc, while it is more likely to find voters with high degrees of partisanship – both towards I and C – in group B). Suppose the groups are equal in size – so there is no reason to target a group because of its size – and that C's possible policies are uniformly distributed over the whole range of ω, i.e., in $\mathbb{C}^M$. Let $\gamma = 0.4$ for both groups. Fig. 2.5 shows the vote share of I (panel a) and the proportion of the electorate that votes for I, C, or abstain (panel b) as functions of ω. There are *two* winning regions for I, with neither of the winning regions containing the center.

The result that a very centrist policy leads to a low vote share for I is easy to explain. Group A, whose members vote fairly uniformly due to concentrated preferences, largely abstains when I is at the center due to positive voting costs. The pro-challenger group B also has a high abstention rate at the center due to positive voting costs, but since they have more dispersed preferences, the proportion of those who vote for C is higher than the proportion of group A that votes for I. Hence, I loses as a centrist due to differential abstention across groups reflecting the dispersion of partisanship within a group. Conceptually, with differential partisan dispersion across groups, if I adopts a position at or very close to $\omega = 0.5$ she does not differentiate herself sufficiently from the challenger, whose expected position is also centered at 0.5, to turn out voters to vote for her.

There are two winning strategies for I. She can favor group A (that is, adopt a platform with ω higher than 0.5) and get the concentrated group A to vote for her heavily while more dispersed group B does not fully vote for C in the same proportions. She can also somewhat favor group B with a position slightly to the left of center, thus getting some support by group B voters whereas voters in the concentrated group A heavily abstain. She needs to run either somewhat to the right or to the left of C.

Several things should be noted. First of all we can no longer classify strategies as targeting swing versus core voters depending on underlying group characteristics. Group A, I's base, has weaker average partisan preferences than group B ($\left|\bar{\lambda}_A\right| < \left|\bar{\lambda}_B\right|$) so that targeting them is consistent with a "mobilize your base" strategy when there are positive voting costs. However, it is also far more concentrated in its preferences, that is, more "swing". As we can see in panel b when comparing votes for I and C, the sharp increase in votes for I as she moves into the high ω winning region reflects the fact that group A voters swing more towards her (albeit from abstention) than group B voters move away as her ω increases.

In the left winning region we see the effect of group A concentration but now a key effect is the sharp shift of votes away from the challenger and towards abstention that allows I to get high vote totals by attracting group B voters. So in this region it is the swing of group A voters towards abstention for $\omega < 0.5$ combined with attracting some group B voters that helps I win.

When C's ω is to the right of center (that is, $\omega \in \mathbb{C}^R$), the winning region to the right of center disappears, and the one to the left of center shifts right, as illustrated in the panel d of Fig. 2.5. This may be explained as follows. If C is right-wing, group A voters are more attracted to her and, for the parameter values given, I can't move far enough right to win them over. The region to the left of center shifts to the right. Some previously winning low ω positions no longer are winners because the group B voters who favor very low ω and were making up part of I's majority before are too dispersed in their partisan preferences to offset the loss of group A voters. On the other side, however, a policy of $\omega = 0.5$ is restored as a winner for I. When C's ω is to the left of center (that is, $\omega \in \mathbb{C}^L$), we see the same two phenomena, but in the opposite direction (panel d of Fig. 2.5), as group B voters who had voted for I now switch to C. Here too the policy of $\omega = 0.5$ is now a winning policy for I.

Consistent perhaps with conventional wisdom, once the challenger moves off-center, centrism for I gets her more votes since it allows her to be enough different from C to induce sufficient turnout to win. However, as we show in section 2.7 below, this apparently intuitive argument is not always correct. It is possible that when the challenger is on one side of the policy spectrum rather than more centrist, I may find it optimal to move in the opposite direction rather than the same direction as C. Furthermore, though it may seem clear that costly voting is necessary for centrism not to work for I, in section 2.6 we will show that when a challenger is popular, it is specifically the fact that voting is costly that may restore centrism as a winning strategy for the incumbent.

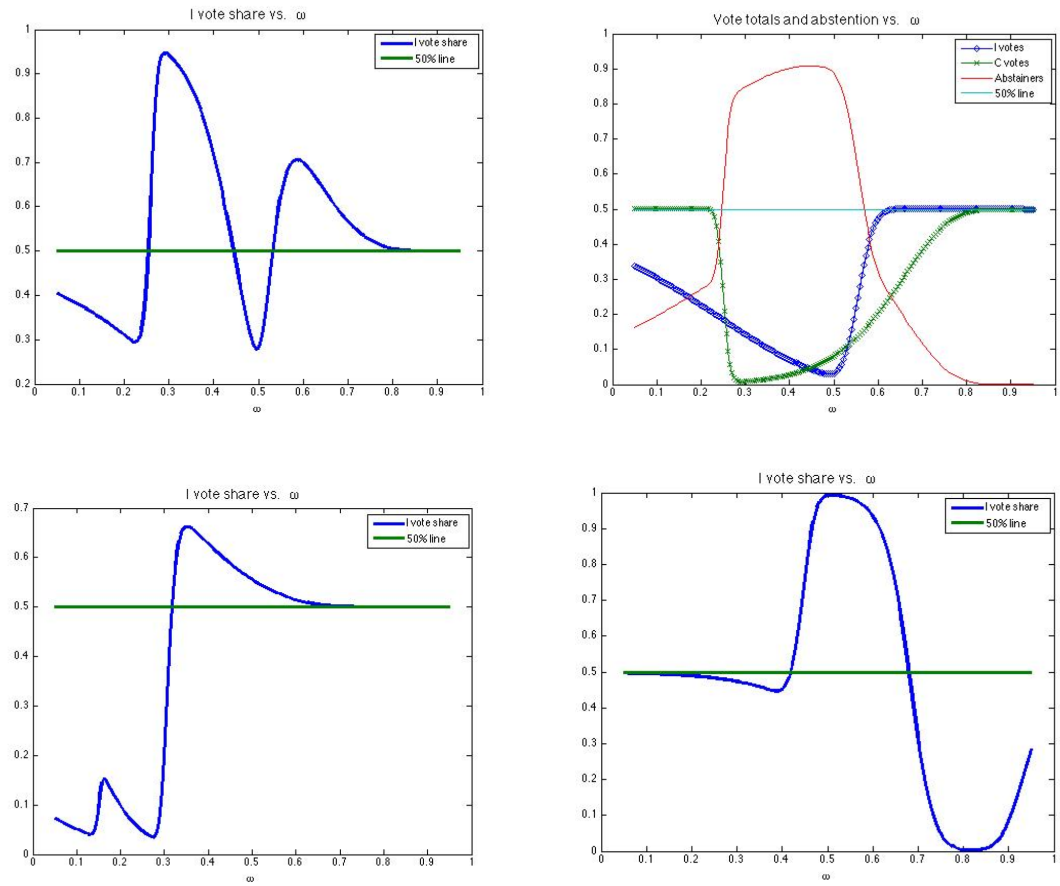

Figure 2.5: Dual Electoral Strategies: (a) vote shares - centrist challenger, (b) vote totals - centrist challenger, (c) vote shares - leftist challenger, (d) vote shares - rightist challenger

2.6 A Popular Challenger

We now consider the case where both groups of voters have a non-policy preference for candidate C. Conventional wisdom is that if a candidate is running against a popular challenger, she is likely to lose the election no matter what electoral strategy she adopts. While this is certainly true if one candidate is popular enough, in this section, we argue that there are winning strategies against a somewhat popular challenger. These winning strategies for I are not necessarily what simple intuition might suggest, and they shed light on some general issues.

2.6.1 Running Away from the Center to Win

We argued that the non-centrist results presented earlier stemmed from the effect of positive voting costs reversing the tendency to adopt centrist policies when voters have concave preferences. In this section we consider another reason why a candidate may choose non-centrism as a winning strategy even under zero voting costs and thus full turnout. To make clear that this is not because we "bias" preferences away from the center, we include a third group of voters with centrist preference, denoted group O. The utility function of the centrists over ω is given by $\frac{\ln \omega + \ln(1-\omega)}{2}$.

Suppose that C is popular, that is, on average he is preferred by voters in both groups A and B on non-policy attributes such as charisma etc. We represent this by $\bar{\lambda}_A = \bar{\lambda}_B = 0.3$. Suppose group A and B each make up 40% of the electorate ($\varphi_A = \varphi_B = .4$), with the remaining 20% being centrists as defined above. We

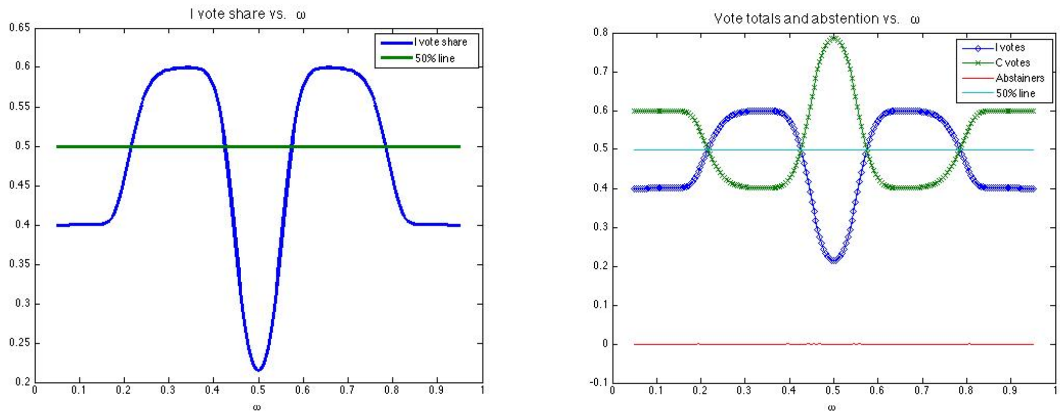

Figure 2.6: Running Away from the Center: (a) vote shares, (b) vote totals

assume that centrists have no average predisposition towards either candidate, that is, $\bar{\lambda}_O = 0$. Assume further that the distribution of candidate preferences in all three voting groups is quite concentrated and identical across the groups, say at $\sigma_A = \sigma_B = \sigma_O = 0.05$. Assume initially that there are no voting costs, i.e., $\gamma = 0$, so that there is full turnout.

Fig. 2.6 plots I's vote share (panel a) and total votes (panel b) as functions of ω when C's policies are neither right- nor left-wing (that is, the challenger's $\omega \in \mathbb{C}^M$). We see that even with no voting costs, centrism does not work because of C's popularity with both groups A and B (panel b around $\omega = 0.5$). I has to favor one group or the other to get their votes – and win with a coalition of that group and centrists – but not so much that she loses the support of the centrist voters.

2.6.2 Targeting a Minority (Voting Costs Restore "Centrism")

When voting costs are positive, centrism may be restored as a vote-getting strategy for I. This is illustrated in Fig. 2.7, with identical parameters to the previous case but with positive voting costs ($\gamma = 0.15$ instead of $\gamma = 0$). A very centrist policy ekes out a bare majority (absent weather shocks). This region is characterized by heavy abstention by group A and group B voters at the center due to sufficiently high voting costs, as can be seen in panel b of Fig. 2.7, combined with a high turnout by centrists that vote for I. She gets more votes than her opponent at $\omega = 0.5$ with less than 20% of the electorate – almost all group O centrist voters – because almost 70% of the electorate that might vote abstain. So voting costs, rather than destroying the strategy of centrism against a centrist opponent, support it.

This result is easy to explain conceptually. I can win because of two key factors. First, most voters abstain because of the similarity of I's policy to her opponent's expected policy. Second, I has the ability to "send a clear message" to those voters for whom $\omega = 0.5$ is the optimal policy inducing them to vote for her. That is, while her opponent also has a likely policy centered on $\omega = 0.5$, concavity of voter preferences means that making clear that policy will be that favored by a specific group beats a message that this is the expected policy. In short, the strategy to beat an ex-ante more popular candidate is to be similar in policy message but "more clear".

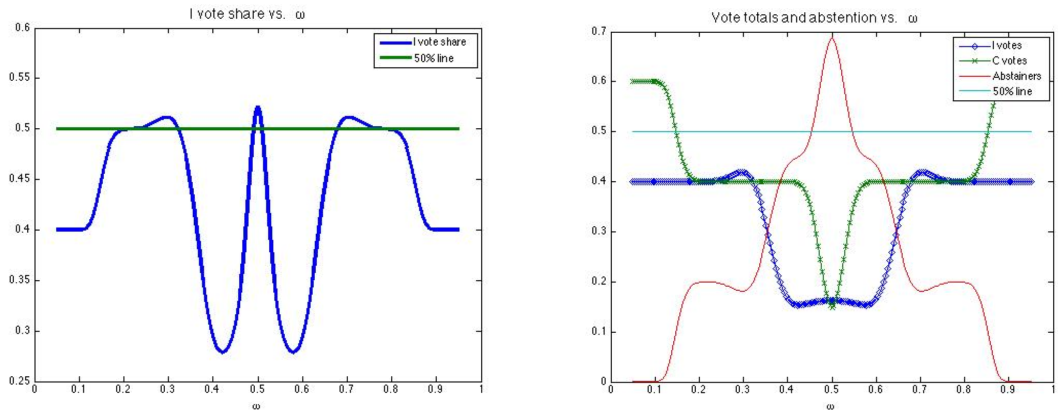

Figure 2.7: Targeting a Minority: (a) vote shares, (b) vote totals

This is not a result about centrism per se, as it could hold at other values of ω when C is believed to have the same average ω and there is likely to be high abstention because voting is costly. A candidate who can send a clearer policy message can win against a popular opponent by targeting the same voting group in terms of promised policy but more "credibly" and count on high abstention from other voters who are not motivated to vote when the two candidates seem similar or the issue is not of sufficient importance to them.

2.7 Targeting Moderate Partisan Voters

2.7.1 Extreme versus Moderate Partisans

The previous section considered targeting minorities by catering exactly to their policy interests while other potential voters abstain. However, if some supporters of a candidate are passionate about her, they will turn out no matter what. Stokes (2005) argues they are so core that they can be taken for granted, as they will always vote for their favored candidate. To investigate the implications of "extreme partisans", we return to the case of two groups of voters, A and B, but suppose that within each group there are both extremists who tend to always vote for one of the candidates independent of the candidate's ω policy and moderates who have less extreme partisan preferences on average and can be swayed by the candidate's policy position. If extremists do not dominate a group and voting costs are low, I will choose her electoral strategy to target a group's moderate partisans.

We now consider a further dimension of differences between groups, the proportion of extremists. Suppose the groups are of equal size but differ in the proportion of extremists. For both types, λ^i is normally distributed around a mean $\bar{\lambda}$, but the mean for extremists is much larger in absolute value than for moderates, so that all (but a tiny number) of them always turn out to vote for their preferred candidate independent of ω.[7] How will differences in the proportion of these extremists with qualitatively different voting behavior affect the electoral strategies that I might adopt? In order to focus on how the existence of extremists might generate multiple vote-getting strategies, we "turn off" the other factors that led to this possibility in sections 2.5.1 and 2.6 by assuming that the groups are equally concentrated ($\sigma_A = \sigma_B$) and that there are no centrist voters.

[7]Since λ^i is normally distributed, a miniscule fraction of "extremists" will have λ^i so low that they will not vote when $\gamma > 0$.

To illustrate how the existence of extreme partisans affects electoral strategies when turnout is crucial, we consider group characteristics such that mobilizing one's base is central to electoral strategies. Remember that in section 2.4.3 (where there were only moderate partisans), the fact that group A was larger but was less likely to vote led I to mobilize group A voters by choice of high ω. This was the only strategy consistent with her winning the election (Fig. 2.3). To represent the problem of candidate I motivating a subset of her base with high voting costs, suppose that group A forms 60% of the electorate ($\varphi_A = .6$) and $\gamma_A = 1$ while $\gamma_B = 0.2$. Assume that extremists in the two groups have mean partisan preferences of $\bar{\lambda}_A = -10$ and $\bar{\lambda}_B = 10$, whereas moderates in the two groups have mean partisan preferences $\bar{\lambda}_A = -0.6$ and $\bar{\lambda}_B = 0.45$. Denote the fraction of extremists in group h by ϵ_h, where we assume that the fraction of extreme partisans is higher in group B, the challenger's base, for example $\epsilon_A = 0.2$ while $\epsilon_B = 0.375$.[8]

Fig. 2.8 shows the vote share of I (panel a), the vote totals for I and C and the fraction who abstain (panel b) as a function of ω. As in the case with only moderate partisans, choosing high ω to mobilize (the moderate part of) her base is a strategy that gains the majority of voters (absent the weather shock). There is however a second strategy consistent with her high vote totals, which is choosing a relatively low ω. Crucial to the existence of this strategy are the extreme partisans in group A who will heavily vote for I independent of her ω. She can combine these

[8]Formally, we simulate this by assuming that the probability that the mean $\bar{\lambda}_h$ of partisan preferences within group h is the extremist value with probability ϵ_h and the moderate value with probability $1 - \epsilon_h$.

voters with group B moderates who have relatively low voting costs and "swing" to her (while group A voters with high voting costs abstain). Hence the existence of voters she can take for granted means that there exist both strategies of mobilizing voters in one's base who might not turn out and of building a "coalition" of extreme partisans and moderates from the other side.

2.7.2 Moving in the Opposite Direction from C

We found in our base cases in section 2.4 that as C's expected policy moves to the right or left, I's high-vote region moves in the same direction. That is, for example, if the challenger is left-wing rather than centrist, i.e., C's ω is in $\mathbb{C}^L$ rather than in $\mathbb{C}^M$, I's high-vote region will also shift to the left. This is intuitive – since her opponent has moved to the left I can as well without endangering her support from voters on the right, that is, group A voters who favor high ω.

We saw the same phenomenon when differences in average partisan preferences and the dispersion of these preferences lead to two winning regions for I when facing a centrist challenger. When C's policy is on one side of the policy spectrum I has only one winning region on the same side of the policy spectrum, and it "moves" in the direction that C has moved. (Compare panels c and d in Fig. 2.5 to panel a as discussed in section 2.5.1.)

This result, as discussed above may seem intuitive – if one's opponent moves to one side of the policy spectrum, a candidate may find it optimal to move more to the center to gain votes. However, the opposite may be true – when C moves away from the center in one direction, I moves away in the opposite direction. To see

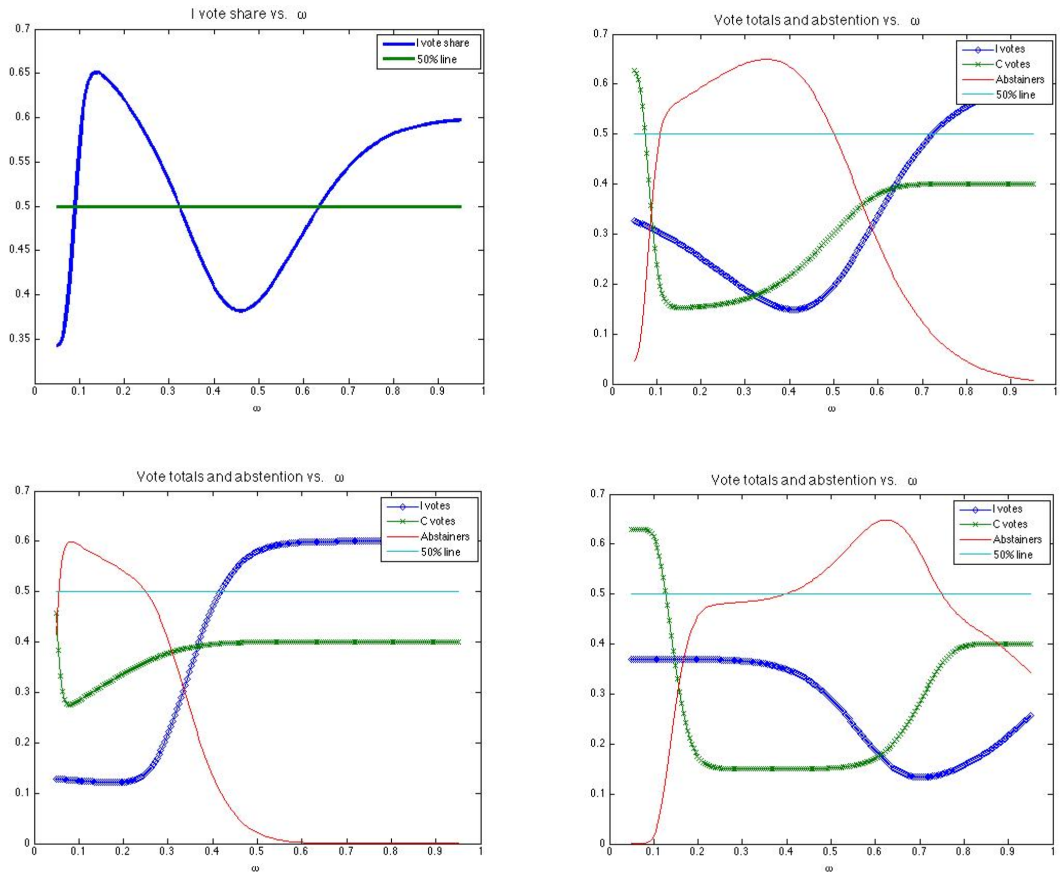

Figure 2.8: Mobilizing Moderates (Taking Extremists for Granted): (a) vote shares - centrist challenger, (b) vote totals - centrist challenger, (c) vote shares - leftist challenger, (d) vote shares - rightist challenger

this, consider the case in which C's ω is in $\mathbb{C}^L$ rather than in $\mathbb{C}^M$. Panel c of Fig. 2.8 shows vote totals of I against a leftist challenger, which can be compared to vote totals against a centrist challenger in panel b of Fig. 2.8. One sees that the winning region of I to the left of center disappears, so that to attract votes I must move to the right. I's winning strategy is now to combine group A moderates with group A extremists. She can no longer win by combining group B moderates with group A extremists (as she could do at the left winning region when the challenger was centrist). That is, whereas she could win against a centrist challenger by swinging moderates not in her base to vote for her, she can no longer do so and must go back to relying on her base by moving in their direction.

Conversely, if the challenger is right-wing rather than centrist, the right-hand winning region for I disappears. (Compare panel d of Fig. 2.8 where C's $\omega \in \mathbb{C}^R$ to panel b where C's $\omega \in \mathbb{C}^M$.) The strategy of relying on her base of group A voters (moderates and extremists) with a high ω is no longer viable and she must instead move left and combine group B moderates with group A extremists if she is to win. In both cases, the intuition is that if a shift in C's position "soaks up" voters in one of I's winning regions, I is then induced to move to the other region in order to win the election. Perhaps this is intuitive, but it does contrast with the intuitive result when there was only one winning region that I's strategy when the challenger moved to one side is to move towards the center, that is, in the same direction. Note further that latter case is often explained by the presence of extreme partisans supporting a candidate allowing her to move more to the center as she can take them for granted. What we see here is that it is their presence that induces the opposite result.

2.7.3 Which Way Does I Move in Response to Changes in C's Position?

As the previous subsection indicates, in theory an exogenous movement by an opponent C could induce I to move in either the same or in the opposite direction. We see cases of both. The rightward movement of the Republicans in 1964 allowed Johnson to position himself more to the center (he probably would have won no matter where he positioned himself) while the leftward movement of the Democrats in the late 1960s allowed Nixon to move more to the center. Conversely, in Britain from the mid 1970's to the late 1980's, as Labor went to the left, the conservatives went to the right.

The latter possibility is not simply a case of a party embracing its traditional base. In the 1950's and early 1960's it is not fully clear whether it was the Democrats or the Republicans who were the party of Civil Rights.[9] The 1964 Civil Rights Act was passed by a coalition of Republicans and Northern Democrats, with geography rather than party affiliation explaining voting behavior (see Enten (2013) for a short summary). The white South was solidly Democratic and solidly against the legislation, while the blacks had traditionally voted Republican. In theory, either party could have moved to the left or right on the issue. By moving right the

[9] We are indebted to Frances Lee for suggesting this example.

Republicans took away the Southern base of the Democrats, who in turn gained the allegiance of black voters.

2.8 Multiple High Vote Regions For I

We now add differential concentration of groups or differential voting costs and find that there can be a greater multiplicity of high vote regions for I.

2.8.1 Differential Concentration of Preferences

Suppose we add differential group concentration to the above case. Consider the parameter configuration in section 2.7 above, but suppose in addition that the groups differ in the concentration of partisan preferences, where group A is far less concentrated in this respect than group B, with $\sigma_A = 0.6$ and $\sigma_B = 0.01$. Group B is extremely concentrated in partisan preferences (and hence votes as a bloc) while group A voters have candidate-specific preferences that are quite dispersed. Fig. 2.9 plots the incumbent's vote share (panel a) and total votes (panel b) as well as the voting behavior of groups A (panel c) and B (panel d) as functions of her policy choice ω for the case where the incumbent faces a challenger whose possible position is uniformly distributed over the whole range $\mathbb{C}^M = [.05, .95]$. There are now three regions in which I's vote total tops 50%.

The emergence of a third high vote region in the center, as well as the change in the shape of the two regions from above reflect the interaction of differential concentration and the other factors in the previous section. As in the case of equal concentration of the two groups in section 2.7, in the left-most region is a combination of group B moderates and group A extreme partisans who supply votes to I, with group A moderates largely abstaining. However, this region becomes a sharp peak (compare panel a in Fig. 2.8 and Fig. 2.9) because of the bloc voting of group B, whose voters respond sharply to marginal changes in I's ω (panel d of Fig. 2.9).

The winning region of high ω found previously also still exists, but it shifts farther to the right and becomes less high. This is also due to the high concentration of partisan preferences in group B and their resultant bloc voting. Interestingly, there is now a third high-vote region in the center between the two regions we observed previously. All moderate B voters abstain and only extremists in that group vote for C while enough group A voters (both moderates and extremists) vote for I so she outpolls her opponent. It is also characterized by sharp changes in vote shares reflecting the high concentration of group B voters.

When the challenger is leftist (Fig. 2.10 panel a) or rightist (Fig. 2.10 panel b) we see the same phenomenon we saw in the case of two winning regions in section 2.7.2 illustrated in Fig. 2.8. Against a leftist challenger I's high vote regions on the left disappear, so she must move right to get above 50% of the vote (by mobilizing her base who prefer high ω). Against a rightist challenger, this "right-wing" strategy is no longer available, so she must adopt a more centrist position.

2.8.2 Heterogeneous Voting Costs

One can get multiple winning regions that look quite similar to those in section 2.8.1 as shown in Fig. 2.9 when heterogeneity within a group comes from voting costs rather than candidate preferences. Suppose that we keep all the parameters

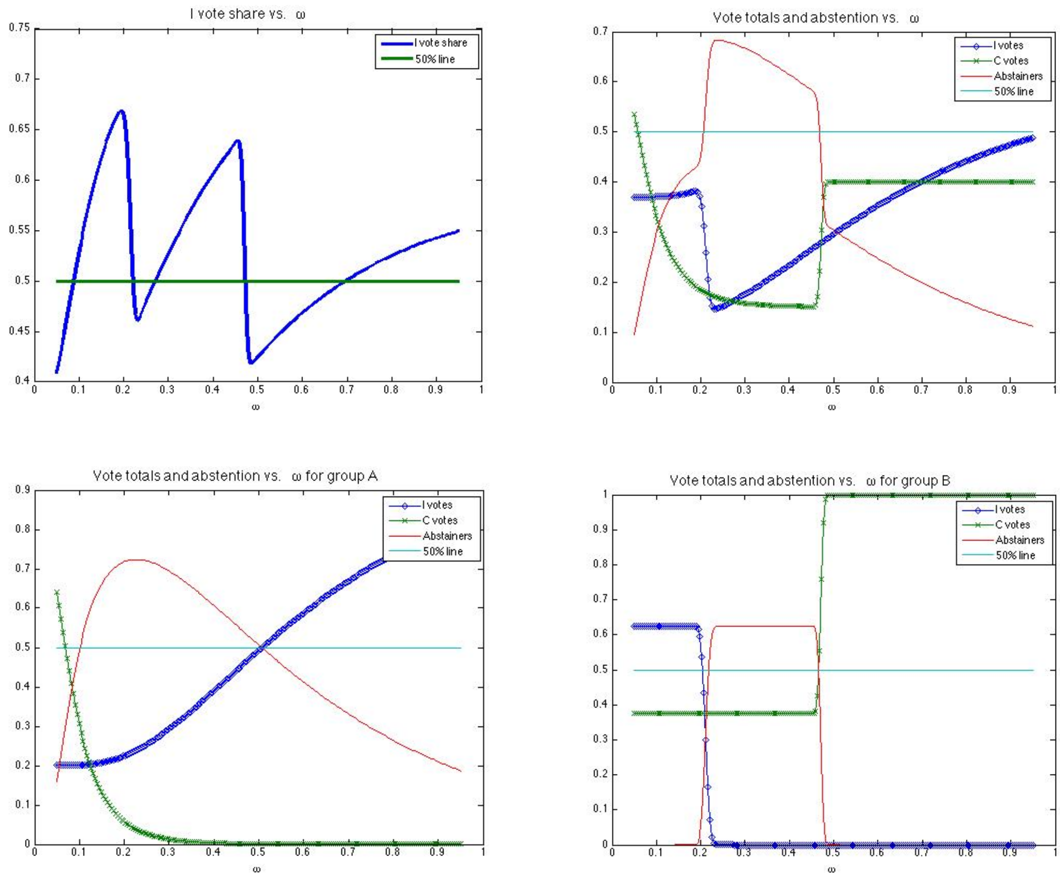

Figure 2.9: Differential Concentration - Centrist Challenger: (a) vote shares, (b) vote totals, (c) vote totals for group A, (d) vote totals for group B

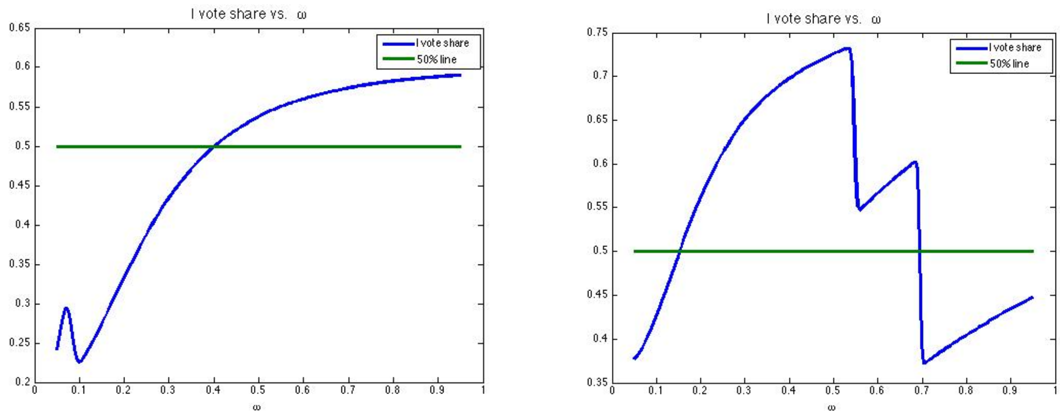

Figure 2.10: Differential Concentration - Non-Centrist Challenger: (a) vote shares - leftist challenger, (b) vote shares - rightist challenger

the same as the above case with three winning regions, except that σ_A and σ_B are the dispersion of voting costs γ rather than of λ within a group. In other words, members within a group differ from each other not in terms of their partisan preferences λ, but instead in the level of their voting costs γ. Fig. 2.11 plots I's vote share and vote totals for this heterogeneous voting cost case.

We see that these graphs are nearly the same as in the previous case with heterogeneous λ but for one difference: around the points of indifference, I's vote share exhibit jumps in the heterogeneous voting cost case (as observed at $\omega = 0.22$) whereas these indifference points result in 50-50 splits of votes between I and C in the heterogeneous λ model, with smooth vote shares for I around them. This arises due to the existence of voters with negative voting costs in the heterogeneous voting cost case. These voters always vote and minor policy changes around indifference points

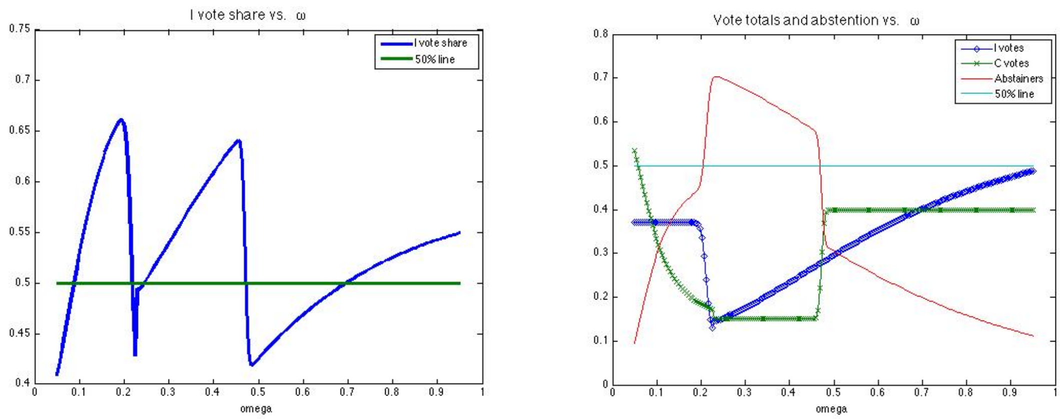

Figure 2.11: Heterogeneous Voting Costs: (a) vote shares - centrist challenger, (b) vote totals - centrist challenger

induces them to switch from voting for I to voting for C (thus skipping abstention), which creates jumps in I's vote share around those points. The magnitudes of these jumps are thus determined by the proportion of voters with negative costs.

2.8.3 Implications of Multiple Winning Regions

The existence of multiple high vote regions are analogous to multiple equilibria for I – in theory she could pick one of several strategies to win an election. What are the implications in practice? Put another way, multiple equilibria imply that the same underlying parameters could be consistent with very different equilibrium choices, but it is not clear that this is what we see in practice. We suggest some reasons for this.

2.8.3.1 Maximizing the Probability of Winning

If there were no exogenous shocks to turnout, so that I knew for sure that all voters who intended to vote actually did (that is, the threshold S were known to be 50% with certainty) then she would be indifferent between any ω that gave her more than 50% of the vote. However, with turnout shocks she would choose platform ω that maximizes her probability of winning, which is simply the value of ω that yields the highest vote share.

2.8.3.2 Most Preferred Ideological Position

Of course, if I preferred some ω over others because of her own ideology, she would choose the winning position most consistent with her ideology. We abstracted from candidate ideology in order to focus on other reasons why candidates take different positions. A ranking of different winning ω with no turnout shocks would yield a unique ω among the winning set analogous to the above argument on maximizing winning probability as the selection mechanism. Introducing both would imply I would choose a position, given the $\mathbb{C}$ of her opponent C that maximizes her expected utility.

2.8.3.3 An Incumbency Disadvantage

A candidate who had to make choices about ω in the past may be tied down by past actions and promises, and hence be able to credibly choose only in a limited part of the ω policy space. Hence, even if multiple winning regions for ω exist in theory, in practice I may be restricted to only one of them. Under this view, it makes sense to think of I as the incumbent, that is, the candidate who is constrained by the fact of previously holding office and having had to make tough choices. We think that in practice it is in fact past history that may select among winning regions.

We further note that the citizen candidate model could be represented as I being restricted to a single ω by voter beliefs. If this value of ω were not in a winning region given what the other candidate (or potential candidates) would do, then I would not run for office.

2.9 Uncertain Voter Preferences

We have assumed so far that I knows the characteristics of groups in choosing her electoral strategies. One may then ask what the effect would be of I being uncertain about these characteristics. We illustrate with uncertainty about a single group characteristic, on the basis of which we can generalize our results.

Suppose the demographic parameters for groups A and B are the same as in section 2.8.1, except that I is uncertain about how partisan are group B moderates, with her assigning a probability $\frac{1}{2}$ to $\bar{\lambda}_B = 0.45$ and $\frac{1}{2}$ to $\bar{\lambda}_B = 0.57$. This is her only uncertainty. A choice of ω is now associated with an expected vote share. We see in Fig. 2.12 two effects of candidate uncertainty on electoral strategies. First, there are now five (expected) high vote regions rather than three. Second, expected vote shares are more sensitive to changes in ω, yielding a sawtooth pattern even sharper than what was observed in Fig. 2.9 when $\bar{\lambda}_B = 0.45$ for group B moderates with certainty.

Intuitively, two possible values for $\bar{\lambda}_B$ for group B moderates is like having three rather than two subgroups in group B, extremists and two types of moderates, each making up $\frac{1}{2}$ of moderate wing. Hence, each of the two high vote regions in the earlier case where moderate group B voters were targeted splits into two expected high vote regions, each half as large in expected value. The more pronounced sawtooth pattern reflects high concentration (i.e., "swingness") of group B moderates.

More generally, uncertainty about parameters is conceptually like having more voting groups, hence more distinct winning regions when these groups are like subgroups of groups which could be targeted to give I a high vote share. Uncertainty about other parameters such as σ_A, σ_B, would have the same conceptual effect.

2.10 Summary and Conclusions

We see this paper as having two main contributions. The first is a model of candidate choice of positions given group characteristics that we think is more realistic than existing models in certain respects, but at the same time highly tractable (and therefore user friendly). The model considers how the ability of a candidate to more precisely indicate her policy positions than her opponent can be used to gain votes. The relative importance of "immutable" candidate characteristics versus "mutable" positions allows the representation of different models of candidate commitment. Unlike most existing models, it allows a central role for abstention in shaping what positions a candidate may adopt. It also presents a simple way to translate choice of positions into probabilities of winning.

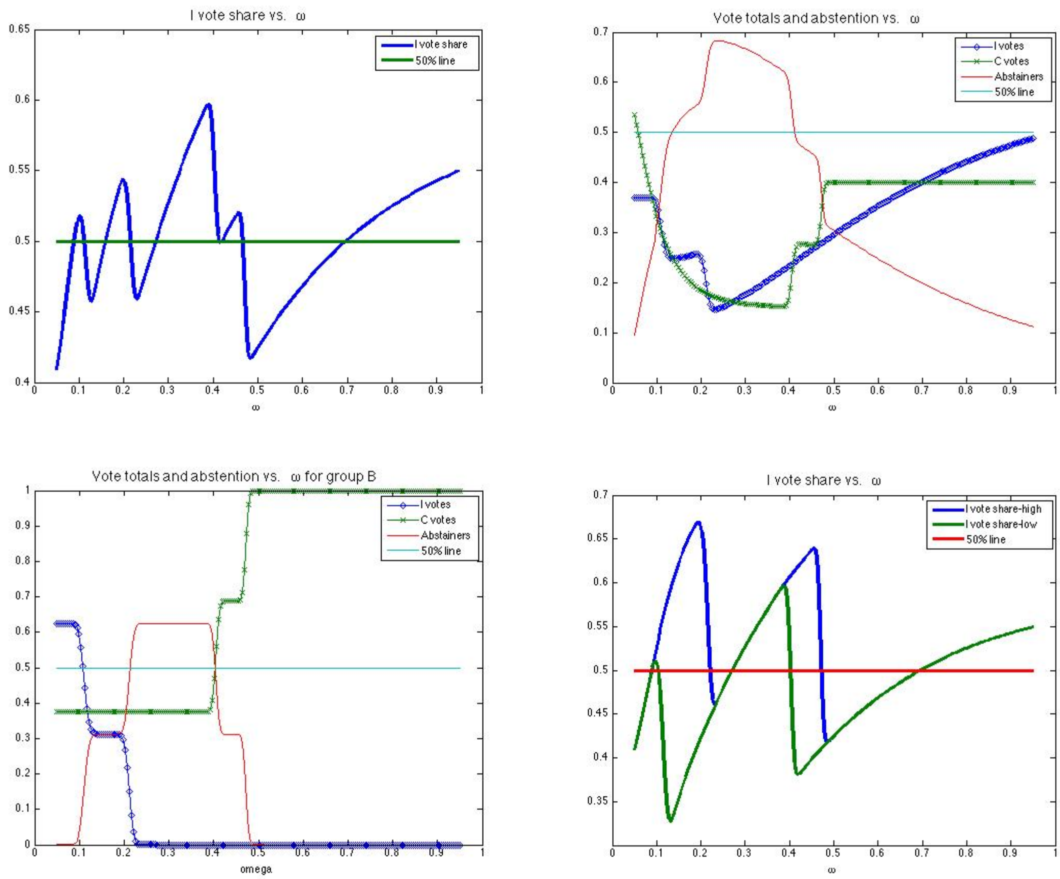

Figure 2.12: Uncertainty on λ_B: (a) expected vote shares, (b) expected vote totals, (c) expected vote totals for group B, (d) range of vote shares

We see the second contribution as the results on strategies a candidate will use to win elections given characteristics of voting groups. The possibility that a candidate can better indicate her position implies that she will often choose a position different than the expected position of her opponent even if the candidate is purely office-motivated and the characteristics of voting groups are known. When voting groups differ primarily in one dimension – for example, concentration of partisan preferences (i.e. "swingness"), intensity of average partisan preferences in the presence of voting costs (likelihood to turn out to vote), we find that standard results on targeting swing versus core groups hold. When they differ in multiple dimensions, the set of electoral strategies becomes richer, sometimes in surprising ways. There may be multiple strategies for attracting a large share of votes where in some cases they cannot be classified either as simply attracting core versus swing voters. When there is only a single strategy, a shift by the opponent to one side of the policy spectrum induces a candidate to become more centrist (to move in the same direction as the opponent has moved), while dual strategies may lead a candidate to adopt the opposite strategy than the opponent when the latter moves to one side of the spectrum (to move in the opposite direction). Though moving to the center is often associated with taking extreme partisans for granted, the presence of such voters – who are sure to vote for one candidate no matter what her position is – may strengthen this latter effect. Finally, in the presence of multiple strategies, an increase in voting costs need not imply targeting one's base by adopting policies they find more favorable than those of the opponent, but adopting the same position as that the opponent is expected to take in order to induce large abstention.

Appendix A: Appendix for Chapter 1

A.1 Expected Payoffs of Voting

Using the structure of B_{ij}'s and the calculus of voting equation, we can write the expected payoffs of voting for each candidate for voters in each group as follows:

For voters in T_1, we have:

$$R_{i1} = p_{i12}y + p_{i13}z + D_i - C_i$$

$$R_{i2} = -p_{i21}y + p_{i23}(z - y) - C_i$$

$$R_{i3} = -p_{i31}z - p_{i32}(z - y) - C_i$$

Note that voting for candidate 3 is never optimal for voters in T_1 since $R_{i3} < R_{i1}$ for voters in T_1. If voters in T_1 were to strategically desert candidate 1, they would vote for candidate 2 (second choice of voters in T_1) and never for candidate 3 (worst outcome for voters in T_1).

For voters in T_2, we have:

$$R_{i1} = -p_{i12}y + p_{i13}(z - 2y) - C_i$$

$$R_{i2} = p_{i21}y + p_{i23}(z - y) + D_i - C_i$$

$$R_{i3} = -p_{i31}(z - 2y) - p_{i32}(z - y) - C_i$$

Note that when $y \leq z/2$, voting for candidate 3 is never optimal for voters in T_2 since $R_{i3} < R_{i2}$ and when $y \geq z/2$, voting for candidate 1 is never optimal for voters in T_2 since $R_{i1} < R_{i2}$ for voters in T_2. This is because the second choice of voters in T_2 is candidate 1 if $y < z/2$ and candidate 3 if $y > z/2$. Voters in T_2 are indifferent between candidates 1 and 3 when $y = z/2$.

For voters in T_3, we have:

$$R_{i1} = -p_{i12}y - p_{i13}z - C_i$$

$$R_{i2} = p_{i21}y - p_{i23}(z - y) - C_i$$

$$R_{i3} = p_{i31}z + p_{i32}(z - y) + D_i - C_i$$

Note that voting for candidate 1 is never optimal for voters in T_3 since $R_{i1} < R_{i3}$ for voters in T_3. Voters in T_3 vote for candidate 2 (their second choice) if they strategically desert candidate 3.

A.2 Proof of Proposition 1

First define the function $g_i(t_1, t_2, t_3) = p_i(t_1, t_2, t_3)y$ for $i = 1, 2, 3$, where $p_i(t_1, t_2, t_3)$ is the binomial probability of being pivotal for voters in T_i when voters in T_i adopt the threshold t_i. Next consider the mapping from $[0, y]^3$ into itself defined by:

$$x_1 = g_1(t_1, t_2, t_3)$$

$$x_2 = g_2(t_1, t_2, t_3)$$

$$x_3 = g_3(t_1, t_2, t_3)$$

Since g_1, g_2 and g_3 are continuous functions and $[0, y]^3$ is compact and convex, by Brouwer's fixed point theorem there exists (x_1^*, x_2^*, x_3^*) such that:

$$x_1^* = g_1(x_1^*, x_2^*, x_3^*)$$

$$x_2^* = g_2(x_1^*, x_2^*, x_3^*)$$

$$x_3^* = g_3(x_1^*, x_2^*, x_3^*)$$

Let $(t_1^*, t_2^*, t_3^*) = (x_1^*, x_2^*, x_3^*)$. Then (t_1^*, t_2^*, t_3^*) is a PBE in $[0, y]^3$.

A.3 Computational Algorithm

Here are the steps of the computational algorithm: I first set the parameter values for $N_1, N_2, N_3, y, \bar{D}, \bar{C}$. Then I start with initial guesses for pivot probabilities p_1^*, p_2^*, p_3^*.[1] These initial guesses imply the thresholds t_1^*, t_2^*, t_3^* as $t_1^* = p_1^* y$, $t_2^* = p_2^* y$ and $t_3^* = p_3^* y$. Next step is to calculate the probability that a voter in T_1 votes, the probability that a voter in T_2 votes, the probability that a voter in T_3 votes sincerely, and the probability that a voter in T_3 votes strategically, which are as follows:

$$P(vote|T_1) = P(C_i - D_i < p_1^* y)$$

$$P(vote|T_2) = P(C_i - D_i < p_2^* y)$$

$$P(sincere|T_3) = P(D_i > p_3^* y \text{ and } D_i > C_i) = P(D_i > p_3^* y).P(C_i - D_i < 0|D_i > p_3^* y)$$

$$P(strategic|T_3) = P(D_i < p_3^* y \text{ and } C_i < p_3^* y) = P(D_i < p_3^* y).P(C_i < p_3^* y)$$

To find these probabilities, I use the formula for the cumulative distribution function of the sum of two random variables that follow uniform distributions $U(a_1, b_1)$ and $U(a_2, b_2)$, which is given as follows:

[1] p_k^* is the probability that a voter in T_k becomes pivotal in making party 1 or party 2 (whichever is preferred) win the election over the other (by either equalizing party 1 votes with party 2 or by breaking the tie between party 1 and party 2 with his vote).

$$F(z) = \begin{cases} 0 & \text{if } z < a_1 + a_2 \\ \frac{(z-a_1-a_2)^2}{2(b_1-a_1)(b_2-a_2)} & \text{if } a_1 + a_2 \leq z < \min\{a_1 + b_2, a_2 + b_1\} \\ \frac{2z-2a_1-b_2-a_2}{2(b_1-a_1)} & \text{if } a_1 + b_2 \leq z < a_2 + b_1 \\ \frac{2z-2a_2-b_1-a_1}{2(b_2-a_2)} & \text{if } a_2 + b_1 \leq z < a_1 + b_2 \\ 1 - \frac{(z-b_1-b_2)^2}{2(b_1-a_1)(b_2-a_2)} & \text{if } \max\{a_1 + b_2, a_2 + b_1\} \leq z < b_1 + b_2 \\ 1 & \text{if } b_1 + b_2 \leq z \end{cases}$$

Using this formula, the probability that a voter in T_1 votes is:

$$P(vote|T_1) = \begin{cases} \frac{p_1^* y}{\bar{C}} + \frac{\bar{D}}{2\bar{C}} & \text{if } p_1^* y \leq \bar{C} - \bar{D} \\ 1 - \frac{(\bar{C}-p_1^* y)^2}{2\bar{C}\bar{D}} & \text{if } \bar{C} - \bar{D} < p_1^* y < \bar{C} \\ 1 & \text{if } p_1^* y \geq \bar{C} \end{cases}$$

The probability that a voter in T_2 votes is:

$$P(vote|T_2) = \begin{cases} \frac{p_2^* y}{\bar{C}} + \frac{\bar{D}}{2\bar{C}} & \text{if } p_2^* y \leq \bar{C} - \bar{D} \\ 1 - \frac{(\bar{C}-p_2^* y)^2}{2\bar{C}\bar{D}} & \text{if } \bar{C} - \bar{D} < p_2^* y < \bar{C} \\ 1 & \text{if } p_2^* y \geq \bar{C} \end{cases}$$

The probability that a voter in T_3 votes sincerely for party 3 is:

$$P(sincere|T_3) = \begin{cases} (\frac{1-p_3^* y}{\bar{D}})(\frac{p_3^* y+\bar{D}}{2\bar{C}}) & \text{if } p_3^* y \leq \bar{D} \leq \bar{C} \\ (\frac{1-p_3^* y}{\bar{D}})(1-\frac{(\bar{C}-p_3^* y)^2}{2\bar{C}(\bar{D}-p_3^* y)}) & \text{if } p_3^* y \leq \bar{C} \leq \bar{D} \\ \frac{1-p_3^* y}{\bar{D}} & \text{if } \bar{C} \leq p_3^* y \leq \bar{D} \\ 0 & \text{otherwise} \end{cases}$$

The probability that a voter in T_3 votes strategically for party 2 is:

$$P(strategic|T_3) = \begin{cases} \frac{p_3^* y}{\bar{C}} \frac{p_3^* y}{\bar{D}} & \text{if } p_3^* y \leq \bar{D} \text{ and } p_3^* y \leq \bar{C} \\ \frac{p_3^* y}{\bar{D}} & \text{if } \bar{C} \leq p_3^* y \leq \bar{D} \\ \frac{p_3^* y}{\bar{C}} & \text{if } \bar{D} \leq p_3^* y \leq \bar{C} \\ 1 & \text{otherwise} \end{cases}$$

Given the above probabilities, the probability that a voter in T_k abstains can be found by subtracting the probability that a voter in T_k votes from 1.[2] Above are the probabilities corresponding to how a single voter votes. Using binomial probability formula, these give the probability that any specified number of voters out of N_1 vote, any specified number of voters out of N_2 vote, any specified number of voters out of N_3 vote sincerely and any specified number of voters out of N_3 vote strategically. Probability of being pivotal for a voter i in T_2 and T_3 is then the summation of probabilities of all the cases that (without i's vote) number of votes for party 2 is equal to or one less than the number of votes for party 1. Probability

[2]The probability that a voter in T_3 votes is the sum of the probability of a sincere vote and the probability of a strategic vote.

of being pivotal for a voter i in T_1 is the summation of probabilities of all the cases that (without i's vote) number of votes for party 1 is equal to or one less than the number of votes for party 2.

Hence, the initial guesses for p_1^*, p_2^*, p_3^* give resulting probabilities of being pivotal. As long as any of the initial guesses is different than the resulting probability, I update the initial guesses slowly towards the resulting probabilities and redo all the steps described above. Equilibrium probabilities p_1^*, p_2^*, p_3^* are found when the initial guesses and the resulting probabilities converge.

A.4 Minority Group Size

I present here another experiment with the pivotal voter model, which is to evaluate the effects of changes in the size of the minority group (N_3) on strategic vs. sincere voting and turnout. Fig.A.1 plots expected fractions of sincere and strategic voters, vote shares of each candidate, participation rates for each group and equilibrium probability of being pivotal as functions of N_3.

There is an inverse U-shaped relationship between the expected fraction of strategic voters and the size of the minority group, as well as between turnout rates and the size of the minority group. This is essentially an optimal response to the inverse U-shaped relationship between the probability of a pivot event and N_3. For low values of N_3, increasing N_3 makes pivot events more likely since it counteracts the difference of strength between party 1 and party 2 and makes the election closer in expectation, thus increases the extent of strategic voting.[3] For high values of N_3, further increasing N_3 means making the election more one-sided towards party 2, in which case there is less need for strategic votes by voters in T_3 in order for party 2 to win the election, which reduces the extent of strategic voting. Thus, minority group size affects strategic voting and turnout through its impact on closeness of the election between front-runners.

A.5 Intermediate Steps of the Solution

Using the fact that party 2 wins the election when the number of votes for party 2 exceeds the number of votes for party 1, we can write $P(\sigma)$ as follows:

$$P(\sigma) = Pr(\tilde{q}_2\mu_2 + \frac{b+1}{2}\mu_3\sigma > \tilde{q}_1\mu_1) = Pr(\tilde{q}_1\frac{2\mu_1}{\mu_3(b+1)} - \tilde{q}_2\frac{2\mu_2}{\mu_3(b+1)} < \sigma)$$

Using $\mu_1 = \mu_2 + x$ and $\mu_1 + \mu_2 + \mu_3 = 1$, we can write μ_1 and μ_2 in terms of μ_3 and x to be $\mu_1 = \frac{1+x-\mu_3}{2}$ and $\mu_2 = \frac{1-x-\mu_3}{2}$. Using these, we have $\tilde{q}_1\frac{2\mu_1}{\mu_3(b+1)} \sim U(\frac{b(1+x-\mu_3)}{(b+1)\mu_3}, \frac{1+x-\mu_3}{(b+1)\mu_3})$ and $-\tilde{q}_2\frac{2\mu_2}{\mu_3(b+1)} \sim U(-\frac{1-x-\mu_3}{(b+1)\mu_3}, -\frac{b(1-x-\mu_3)}{(b+1)\mu_3})$. Then, $P(\sigma)$ can be figured out from the cumulative distribution function of the sum of two independent uniform random variables. Specifically, $P(\sigma) = F(\sigma)$ where F(.) is the cumulative distribution function of $\tilde{q}_1\frac{2\mu_1}{\mu_3(b+1)} - \tilde{q}_2\frac{2\mu_2}{\mu_3(b+1)}$. The derivative of $P(\sigma)$ is equal to $P'(\sigma) = f(\sigma)$ where $f(.)$ is the probability distribution function of $\tilde{q}_1\frac{2\mu_1}{\mu_3(b+1)} - \tilde{q}_2\frac{2\mu_2}{\mu_3(b+1)}$.

[3]This logic applies when $x > 0$, i.e. when party 1 has more supporters than party 2. When $x < 0$, expected fraction of strategic voters is monotonically decreasing in N_3.

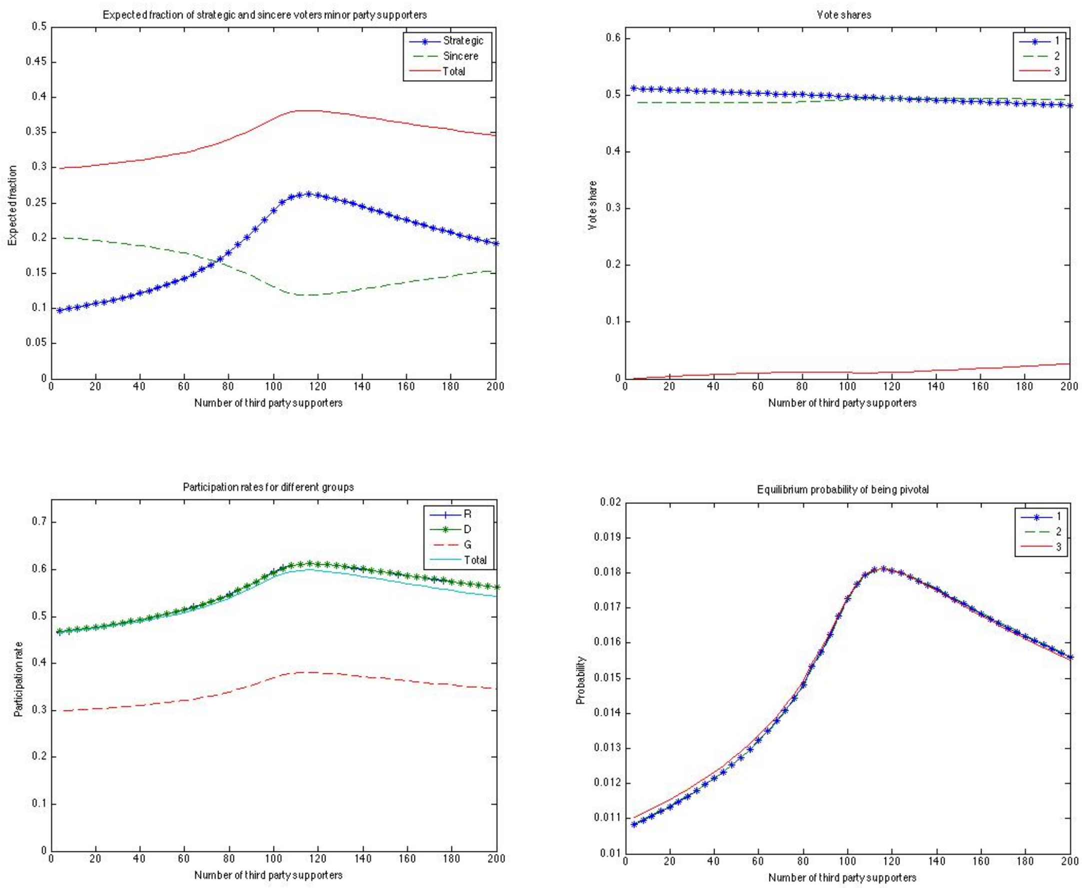

Figure A.1: Third Party Voting vs. Number of Third Party Supporters: (a) sincere and strategic voting, (b) vote shares (c) participation rates, (d) pivot probabilities

The first order condition of this maximization problem is:

$$yP'(\sigma) - \bar{D}\sigma = yf(\sigma) - \bar{D}\sigma \begin{cases} \leq 0 & \text{if } \sigma = 0 \\ = 0 & \text{if } \sigma \in (0,1) \\ \geq 0 & \text{if } \sigma = 1 \end{cases}$$

In order to solve this maximization problem, $P(\sigma) = F(\sigma)$ needs to be figured out. To do that, we need to use the formula of $F(.)$, which is the cumulative distribution function of the sum of two independent uniform random variables. I have already provided the formula for the cumulative distribution function of the sum of two random variables that follow uniform distributions $U(a_1, b_1)$ and $U(a_2, b_2)$ in Appendix A.3. Since we are interested in finding the probability $Pr(\tilde{q}_1 \frac{2\mu_1}{\mu_3(b+1)} - \tilde{q}_2 \frac{2\mu_2}{\mu_3(b+1)} < \sigma)$, we have $a_1 = \frac{b(1+x-\mu_3)}{(b+1)\mu_3}, b_1 = \frac{1+x-\mu_3}{(b+1)\mu_3}, a_2 = -\frac{1-x-\mu_3}{(b+1)\mu_3}, b_2 = -\frac{b(1-x-\mu_3)}{(b+1)\mu_3}, z = \sigma$ for this case.

A.6 Proofs of Propositions 2-4

Looking at the expression for the cutoff σ^*, if the absolute value of x is low enough to induce a positive extent of strategic voting (when $\sigma^* > 0$), two cases emerge: When $\frac{y}{\bar{D}} < \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$, we have that $\frac{d\sigma^*}{dy} \geq 0$, $\frac{d\sigma^*}{d\mu_3} \geq 0$, $\frac{d\sigma^*}{d\bar{D}} \leq 0$, $\frac{d\sigma^*}{dx} \leq 0$. When $\frac{y}{\bar{D}} \geq \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$, we have that $\frac{d\sigma^*}{dy} \geq 0$, $\frac{d\sigma^*}{dx} \geq 0$, $\frac{d\sigma^*}{d\bar{D}} \leq 0$. Also observe that the case $\frac{y}{\bar{D}} < \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$ gets more likely when $\bar{D}$ and x are higher, also when y and μ_3 are lower.

Given these, observe that $\frac{d\sigma^*}{dy} \geq 0$ and $\frac{d\sigma^*}{d\bar{D}} \leq 0$ for both cases (which proves Proposition 2).

Starting from low x (so that $\frac{y}{\bar{D}} \geq \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$), $\frac{d\sigma^*}{dx} \geq 0$ until x reaches a threshold $\tilde{x}$, after which the case $\frac{y}{\bar{D}} < \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$ applies so that $\frac{d\sigma^*}{dx} \leq 0$ (which proves Proposition 3).

Starting from low μ_3 (so that $\frac{y}{\bar{D}} < \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$), $\frac{d\sigma^*}{d\mu_3} \geq 0$ until μ_3 reaches a threshold $\tilde{\mu_3}$, after which the case $\frac{y}{\bar{D}} \geq \frac{(1-b)2x(1+x-\mu_3)}{(1+b)^2\mu_3^2}$ applies so that the sign of the derivative $\frac{d\sigma^*}{d\mu_3}$ is ambiguous (which proves Proposition 4).[4]

A.7 Finite Voter Version and the Size Result

I here present the finite voter version of the ethical voter model. Setup is almost the same as the ethical voter model with continuum of voters, the only differences are that the number of voters is finite and each voter votes with probability π instead of total turnout rates being stochastic.[5]

I present two experiments with different parameter values: The first experiment is to multiply the base values of $N_1 = 100, N_2 = 99, N_3 = 2$ with the scalars $1, 2, ..., 10$ to see the impact of electorate size on strategic voting choice while keeping the shares of each group (μ_1, μ_2, μ_3) the same. Other parameter values are set

[4] $\tilde{x}$ and $\tilde{\mu_3}$ can be found by solving $\frac{y}{\bar{D}} = \frac{(1-b)2\tilde{x}(1+\tilde{x}-\mu_3)}{(1+b)^2\mu_3^2}$ and $\frac{y}{\bar{D}} = \frac{(1-b)2x(1+x-\tilde{\mu_3})}{(1+b)^2\tilde{\mu_3}^2}$.

[5] The assumption that each voter votes with probability π is the same as the one in the pivotal voter model with finite voters. Due to the finite number of voters, I make use of computational methods to solve this version of ethical voter model.

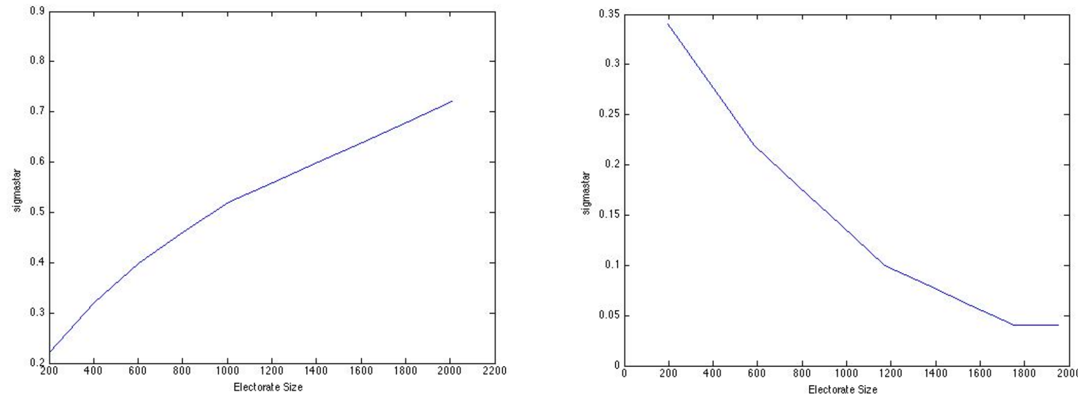

Figure A.2: Third Party Voting vs. Electorate Size - Rule-Utilitarian Model: (a) Base values: $N_1 = 100, N_2 = 99, N_3 = 2$, (b) Base values: $N_1 = 90, N_2 = 100, N_3 = 5$

to be $y = 4, \bar{D} = 2, \pi = 0.6$. For the second experiment, I change the base values to $N_1 = 90, N_2 = 100, N_3 = 5$ and redo the same experiment.

Fig.A.2 plots the optimal extent of strategic voting (σ^*) as a function of electorate size for both experiments. Panel a shows that the extent of strategic voting is increasing (hence third party vote share is decreasing) in electorate size for the first experiment whereas panel b shows that the extent of strategic voting is decreasing (hence third party vote share is increasing) in electorate size for the second experiment.

This is essentially a law of large numbers result:[6] Since every voter's act of voting vs. abstaining is a draw from the Bernoulli distribution, increasing the electorate size is essentially increasing the number of draws from the distribution. Therefore, as the electorate gets larger, election results converge to the true data generating process (the election outcome gets more likely to be determined by relative strength of parties rather than turnout shocks). In cases like first (where μ_1 is close to μ_2, thus increasing electorate size makes it more likely that the election is close), this makes strategic voting more likely to affect the election outcome, thus increases the extent of strategic voting. In cases like second (where the difference between μ_1 and μ_2 is greater, thus increasing electorate size makes it more likely that the election is one-sided towards party 2), this makes strategic voting less likely to affect the election outcome, thus decreases the extent of strategic voting.

Hence, for the ethical voter model, the relationship between the extent of strategic voting (thus the third party vote share) and electorate size is in strong interaction with the relative size of groups. Unlike the pivotal voter model, ethical voter model does not give clear-cut predictions on the effects of electorate size.

[6] For a complete statement of the two forms of law of large numbers (Khinchine's weak law of large numbers and Kolmogorov's strong law of large numbers), see any graduate level statistics or econometrics textbook. Greene (2011) for example, provides one.

Bibliography

[1] Bensel, R. F., & Sanders, M. E. (1979). The effect of electoral rules on voting behavior: the electoral college and shift voting. *Public Choice, 34*(1), 69-85.

[2] Besley, T., & Coate, S. (1997). An economic model of representative democracy. *The Quarterly Journal of Economics, 112*(1), 85-114.

[3] Brender, A., & Drazen, A. (2013). Elections, leaders, and the composition of government spending. *Journal of Public Economics, 97*, 18-31.

[4] Brennan, G., & Buchanan, J. (1984). Voter choice: evaluating political alternatives. *American Behavioral Scientist, 28*(2), 185-201.

[5] Brennan, G., & Hamlin, A. (1998). Expressive voting and electoral equilibrium. *Public Choice, 95*(1), 149-175.

[6] Brennan, G., & Lomasky, L. (1997). *Democracy and decision: The pure theory of electoral preference.* Cambridge: Cambridge University Press.

[7] Burden, B. C. (2005). Minor parties and strategic voting in recent US presidential elections. *Electoral Studies, 24*(4), 603-618.

[8] Callander, S. (2005). Electoral competition in heterogeneous districts. *Journal of Political Economy, 113*(5), 1116-1145.

[9] Castanheira, M. (2003). Why vote for losers?. *Journal of the European Economic Association, 1*(5), 1207-1238.

[10] Coate, S., & Conlin, M. (2004). A group rule-utilitarian approach to voter turnout: theory and evidence. *American Economic Review, 94*(5), 1476-1504.

[11] Cox, G. W. (1994). Strategic voting equilibria under the single nontransferable vote. *American Political Science Review, 88*(3), 608-621.

[12] Cox, G.W. (2009). Swing voters, core voters, and distributive politics. In I. Shapiro, S. Stokes, E.J. Wood & A. Kirshner (Eds.), *Political Representation*. Cambridge: Cambridge University Press.

[13] Cox, G. W., & McCubbins, M. D. (1986). Electoral politics as a redistributive game. *The Journal of Politics, 48*(2), 370-389.

[14] Dixit, A., & Londregan, J. (1996). The determinants of success of special interests in redistributive politics. *The Journal of Politics, 58*(4), 1132-1155.

[15] Drazen, A., & Eslava, M. (2010). Electoral manipulation via voter-friendly spending: theory and evidence. *Journal of Development Economics, 92*(1), 39-52.

[16] Dube, A., Lester, T. W., & Reich, M. (2010). Minimum wage effects across state borders: estimates using contiguous counties. *The Review of Economics and Statistics, 92*(4), 945-964.

[17] Duverger, M. (1954). *Political parties, their organization and activity in the modern state.* New York: Wiley.

[18] Edlin, A., Gelman, A., & Kaplan, N. (2007). Voting as a rational choice: why and how people vote to improve the well-being of others. *Rationality and Society, 19*(3), 293-314.

[19] Enten, H. (2013, August 28). Were Republicans really the party of civil rights in the 1960s?. *The Guardian*. Retrieved from http://www.theguardian.com/commentisfree/2013/aug/28/republicans-party-of-civil-rights.

[20] Feddersen, T., & Sandroni, A. (2006). A theory of participation in elections. *American Economic Review, 96*(4), 1271-1282.

[21] Greene, W. H. (2011). *Econometric analysis.* Upper Saddle River, NJ: Prentice Hall.

[22] Hamlin, A., & Jennings, C. (2011). Expressive political behaviour: foundations, scope and implications. *British Journal of Political Science, 41*(3), 645-670.

[23] Jankowski, R. (2002). Buying a lottery ticket to help the poor: altruism, civic duty, and self-interest in the decision to vote. *Rationality and Society, 14*(1), 55-77.

[24] Kahneman, D., & Tversky, A. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.

[25] Kawai, K., & Watanabe, Y. (2013). Inferring strategic voting. *American Economic Review, 103*(2), 624-662.

[26] Krasa, S., & Polborn, M. K. (2012). Political competition between differentiated candidates. *Games and Economic Behavior, 76*(1), 249-271.

[27] Lacy, D., & Burden, B. C. (1999). The vote-stealing and turnout effects of Ross Perot in the 1992 US presidential election. *American Journal of Political Science, 43*(1), 233-255.

[28] Ledyard, J. O. (1984). The pure theory of large two-candidate elections. *Public Choice, 44*(1), 7-41.

[29] Lindbeck, A., & Weibull, J.W. (1987). Balanced budget redistribution as the outcome of political competition. *Public Choice, 52*(3), 273-97.

[30] Mackie, G. (2011). *An examination of the expressive theory of voting*. Unpublished manuscript, University of California, San Diego.

[31] Matakos, K., & Xefteris, D. (2017). Divide and rule: redistribution in a model with differentiated candidates. *Economic Theory, 63*(4), 867-902.

[32] Meirowitz, A. (2005). Informational party primaries and strategic ambiguity. *Journal of Theoretical Politics, 17*(1), 107-136.

[33] Myatt, D. (2015). *A theory of voter turnout*. Working paper.

[34] Myerson, R. B., & Weber, R. J. (1993). A theory of voting equilibria. *American Political Science Review, 87*(1), 102-114.

[35] Osborne, M. J., & Slivinski, A. (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics, 111*(1), 65-96.

[36] Palfrey, T. R. (1984). Spatial equilibrium with entry. *The Review of Economic Studies, 51*(1), 139-156.

[37] Palfrey, T. R., & Rosenthal, H. (1985). Voter participation and strategic uncertainty. *American Political Science Review, 79*(1), 62-78.

[38] Roemer, J. E. (1998). Why the poor do not expropriate the rich: an old argument in new garb. *Journal of Public Economics, 70*(3), 399-424.

[39] Spenkuch, J. (2013). *On the extent of strategic voting*. Working paper.

[40] Stokes, S. C. (2005). Perverse accountability: a formal model of machine politics with evidence from Argentina. *American Political Science Review, 99*(3), 315-325.