

# TACTICAL COORDINATION IN PLURALITY ELECTORAL SYSTEMS

DAVID P. MYATT  
STEPHEN D. FISHER  
*University of Oxford*<sup>1</sup>

*Simple plurality election systems (commonly known as 'first past the post') are often associated with the dominance of two political parties. Such systems tend to reward leading parties with a disproportionately large number of seats (the 'mechanical' effect) and provoke tactical voting, where voters switch away from trailing parties (the 'psychological' effect). We view tactical voting as a coordination problem. A group of voters wish to prevent a win by a disliked party (such as the Conservatives in recent UK elections) and must partially coordinate behind a single challenger (such as Labour or the Liberal Democrats) in order to do this. Crucially, voters have limited information on the situation within their constituency and hence there is no common knowledge of the game being played—tactical voting is a global game. We show that in this setting, voters will only partially coordinate. Furthermore, tactical voting exhibits negative feedback—tactical voting by others reduces the incentive for an individual to vote tactically, since they become concerned that they may switch in the wrong direction. We calibrate our model, and apply it to the UK General Election of 1997. Throughout England, we find that the 'mechanical' and 'psychological' effects tend to offset each other: tactical voting serves to reverse the Conservative bias that results from the geographic distribution of preferences.*

## I. DUVERGER'S LAW AND TWO-PARTY SYSTEMS

Political scientists (perhaps in common with economists) are not known for their ability to reach a consensus. Nevertheless, they generally agree that

simple plurality (or first past the post (FPTP)) electoral systems tend to be associated with the dominance of two main political parties.<sup>2</sup> This proposition is known as Duverger's Law, but the origins of the idea are older than his contributions.<sup>3</sup> Duverger (1954) argued that the association was causal, so

<sup>1</sup> The authors thank Hyun Shin, Chris Wallace, and the referees for helpful comments and suggestions.

<sup>2</sup> The plurality system of the United Kingdom works as follows. The country is divided into single member districts or constituencies. Each individual casts a single vote for one candidate. The candidate with the largest number of votes (a plurality) is declared the winner. In the event of a tie, the returning officer holds a casting vote. By convention, the office will resolve the tie with the toss of a fair coin.

<sup>3</sup> Riker (1982) provides a historical account of the law's development within political science.

that when a state adopts a simple plurality electoral system the effective number of political parties will converge towards a political duopoly. Likewise, he hypothesized that a change from a simple plurality to a more proportional system would allow the number of effective parties to increase. By way of illustration, Duverger compared the electoral history of Belgium to that of the United Kingdom at the beginning of the twentieth century. At the start of the century both countries had a simple plurality electoral system, a Liberal Party in decline, and a rapidly growing socialist party. The next 50 years of British political history saw a switch from a two-party system dominated by the Liberals and Conservatives (or 'Tories') to one involving Labour and the Conservatives. In contrast, a change in the Belgian electoral system from FPTP to proportional representation (PR) appears to have halted the decline in the Liberal Party, permitting the development of a three-party system over the same time period.

Despite the apparent simplicity of this example, there have been persistent doubts over the causal processes driving Duverger's Law. Perhaps the choice of the electoral system reflects political and social divisions within society rather than there being any effect of the system *per se*. Perhaps the existence of other political institutions, such as a presidency, or other features of the political system can explain the observed regularity between the number of parties and the electoral system. The comparative literature on democracies suggests not (Lijphart, 1994; Amorim Neto and Cox, 1997). But such a rejection of alternative explanations is not enough to generate a positive understanding of how the law might work.

Duverger described two processes: a *mechanical effect* and a *psychological effect*. The mechanical effect is generated by the translation of votes into seats under the electoral system. It is common to find that parties with a large share of the vote receive disproportionately large shares of the seats

in the legislature. It follows that seats in the legislature will tend to be concentrated among fewer parties than votes in the election. The extent to which this is the case depends on the geographical distribution of the votes. If there were no variation from district to district, the largest party would win all the seats. Minor parties win more seats when their votes are geographically concentrated.<sup>4</sup>

The second part of the process is the psychological effect. Voters know that only the winning candidate in their district is elected and therefore are mindful of 'wasting' their vote on a candidate who looks to have little chance of winning. Sometimes voters would rather abandon their preferred candidate to vote for another with a better chance of winning so as best to influence the outcome of the election. Riker (1982) cites Henry Droop's (1871) early exposition of this logic:

As success depends upon obtaining a majority of the aggregate votes of all the electors, an election is usually reduced to a contest between the two most popular candidates . . . even if other candidates go to the poll, the electors usually find out that their votes will be thrown away, unless given in favour of one or other of the parties between whom the election really lies.

This phenomenon is known as *tactical voting* in the United Kingdom and *strategic voting* in the United States (Fisher, 2001). The election outcome will be fully Duvergerian when only the two leading candidates receive votes. It is the phenomenon of tactical voting, as the central component of Duverger's psychological effect, that is of interest to us. In particular, we wish to establish whether a coherent theory of instrumental voting (where voters care predominantly about the identity of the winning candidate) will predict strict bipartism or multi-candidate support.

To build such a theory, we recognize that tactical voting arises from the game played by voters in a plurality election. Under such a system, the outcome depends upon the number of votes cast for each

<sup>4</sup> In recent UK elections, the Liberal Democrats (and previously the Liberal/SDP Alliance) have received very few parliamentary seats despite capturing a significant share of the popular vote. There are 529 parliamentary constituencies in England. In 1987, the Liberal/SDP Alliance won only 10 English seats from 23.8 per cent of the popular vote. Their Liberal Democrat successors saw their vote share decline to 17.9 per cent in 1997, and yet captured 34 English seats. A possible explanation is increased geographic concentration. The standard deviation of vote share across constituencies increased from 8.05 to 10.61 over this period, reflecting an increased focus of Liberal votes in areas such as the south-west. Other explanations for the increased electoral success include better coordination of tactical voting, which is, of course, central to the analysis of this paper.

candidate, with the most popular candidate declared the winner. But this, of course, means that the optimal voting decision for an individual will depend on the voting strategies that others are expected to adopt. In fact, we can loosely regard the election as presenting a coordination problem. If a voter cares about the identity of the winner, then she will wish to cast her vote for one of the two leading candidates, for only then can her vote have any effect.<sup>5</sup> The voters successfully coordinate when they all focus their attention on the same pair of candidates. When evaluating the theoretical foundations for Duverger's psychological effect, therefore, we are really asking when the voters will achieve such coordination, and when coordination will fail.

Recent formal theories of tactical voting predict strict bipartism. Palfrey (1989), Cox (1984, 1994, 1997) and Myerson and Weber (1993) all suppose that voters have common and precise knowledge of the distribution of political preferences throughout an electorate. Caring only about the outcome of the election, some supporters of lower-placed candidates switch their votes to one of the two leading contenders. This increases the support of the leaders, while further eroding the support of trailing candidates, further enhancing the incentive to vote tactically. Such a 'bandwagon' (Simon, 1954) process continues indefinitely, and leads to a strictly 'Duvergerian' equilibrium with support for only two candidates. This means that, according to these theoretical analyses, the psychological effect is sufficient for a two-party system, at least at the level of a voting district. Of course, there is still room for the existence of multiple parties at a national level, since different pairs of parties may be dominant in different electoral districts.

Unfortunately, this theoretical prediction does not match empirical observation. Consider, for instance, parliamentary elections in the United Kingdom. Members of the House of Commons are elected by the plurality rule, and the twentieth century saw the dominance of the Labour and Conservative parties—the Duvergerian mechanical effect was in operation. The Liberal Democrats (and before them, the Liberal

Party and the SDP) have traditionally won a relatively small number of seats. Nevertheless, at the constituency level there remains extensive three-party competition, and it is not the case that the vote shares of the third party are reduced to zero. Duverger's psychological effect does not lead all the way to strict local bipartism, and thus the predictions of existing formal game-theoretic models are rejected. This leads us to reconsider the assumptions from which these theoretical predictions are derived.

Formal theories of voting typically assume that voters are instrumentally rational. This means that they care only about influencing the outcome of the election, rather than about the nature of their vote *per se*. The observation of multi-candidate support might lead us to reject this notion of rationality. Before doing so, however, it is worthwhile returning to the informational assumptions that are also employed. We have observed elsewhere (Myatt and Fisher, 2002) that the leading formal theories all assume that every voter shares the same opinion of the true levels of support for each of the competing candidates. In other words, they all understand exactly which game they are playing. In reality, of course, voters have no such common knowledge. Whereas the political climate at the national level may be ascertained by widespread media coverage, including the ever-pervasive opinion polls, the situation at the constituency level is much less clear. Indeed, survey research shows that voters may hold radically different opinions of the 'state of play' in their constituency.<sup>6</sup> When casting his vote, an individual must take account of the fact that other voters may view the constituency situation in a way that is different from his own, and hence contemplate the different voting games that might be in play.

This logic suggests that we should model plurality elections as a global game (see the Assessment in this issue; Carlsson and van Damme, 1993; Morris and Shin, 2001). Employing such an approach leads to some surprising insights. As we explain in this paper, when voters are unsure of the opinions held by others, they become less willing to vote tactically. Very roughly, voters become concerned that they

<sup>5</sup> We recognize that an election may involve a tie for the lead between three or more candidates. We omit this consideration for simplicity.

<sup>6</sup> The 1997 British Election Campaign Study (Norris *et al.*, 1999) shows that 38 per cent of those whose preferred party came third had expected it to come first or second. In contrast, only 13 per cent of those whose preferred party came third had expected another party to win. This would appear to suggest that voters have inflated expectations for their favourite candidate. However, section IV provides good reason to expect such distortions in voter expectations.

might follow a tactical voting bandwagon in the wrong direction—the suspicion that their own information sources may differ from those of others generates caution. This leads to voting equilibria with limited tactical voting, and the failure of voters fully to coordinate behind the leading candidates. The political scientific implications are important. First, the existence of multi-candidate support in equilibrium means that election outcomes are not necessarily inconsistent with instrumental motivations on the part of voters. Second, the resulting vote shares for the candidates will change in response to parameter shifts. In other words, the approach leads to comparative static predictions that may be taken to the data. Earlier theories could not offer such comparative statics, since they predicted pure bipartism. Third, the approach places heavy emphasis on the nature of information sources available to voters, and hence gives new insights into the interplay between social communication and electoral performance.

It is our goal in the remainder of this paper to explain and illustrate this approach. Motivated by recent British General Elections, we build a simple model in which a group of voters wish to coordinate tactically in order to prevent the win of a disliked candidate. Turning to the information sources of voters, we employ a simple framework in which voters become aware of the constituency state of play by observing others within their own communities. We are able to calibrate our model, and apply it to a subset of British constituencies in the 1997 General Election. Doing so, we are able to obtain an estimate of the number of seats lost by the Conservative Party owing to tactical voting. We find that while Duverger's psychological effect is more muted than that predicted by existing formal theories, it none the less had an impact on a significant number of parliamentary seats, and in fact served to offset much of the mechanical effect.

Our paper is structured as follows. In section II we explain the tactical coordination problem faced by a group voters in many elections. We turn to the individual's decision in section III, and argue that anticipated

tactical voting by others will reduce his own tactical incentives. We develop a fuller model in section IV, moving on to calibrate and apply it in section V.

## II. THE TACTICAL COORDINATION PROBLEM

We focus on the heart of the tactical voting problem by restricting to a class of three candidate elections. We suppose that one candidate (perhaps an unpopular incumbent) enjoys some unwavering support. The remaining (perhaps anti-incumbent) voters wish to ensure the defeat of this candidate, and are faced with the choice of two challengers. They may have differing preferences over these challengers, but are willing to switch their vote away from their first choice in order to avoid a win by the disliked third candidate. We find it helpful to review examples of this idea.

The 1997 General Election was a turning point in British political history. After 18 years in power, John Major's incumbent administration was unpopular, and many voters were eager to ensure a Conservative defeat. In all but two of the 529 English constituencies, the election was a three-horse race.<sup>7</sup> Anti-Conservative voters were faced with a choice between Labour and Liberal Democrat candidates. Survey evidence reinforces the common view that most Labour and Liberal Democrat supporters ranked the Conservatives last. Unless their party allegiance was particularly strong, such voters may have been willing to switch their vote between the two opposition parties in order to stop a Conservative win.<sup>8</sup> They were playing a game of 'defeat the Tory.'

The Folkestone and Hythe seat provides a classic illustration of a potential coordination problem faced by such voters. This seat was, and still is, held by the Conservative Michael Howard, who was the Home Secretary prior to the election. He went on to retain his seat, receiving 39 per cent of the popular vote (Table 1(a)). His win was not inevitable. The anti-Tory vote was split almost equally between the

<sup>7</sup> There were two exceptions. In Tatton the Liberal Democrat and Labour candidates withdrew in favour of Martin Bell, who ran on an anti-sleaze ticket against Neil Hamilton. West Bromwich West was Betty Boothroyd's seat. As the Speaker of the House of Commons, she (by convention) faced no opposition from other major party candidates.

<sup>8</sup> We are, of course, neglecting any possible tactical voting by supporters of the Conservatives. The Conservative Party rarely ranks third, and hence there are few opportunities for its supporters to cast a meaningful tactical vote.

**Table 1**  
**Success and Failure of Tactical Coordination in Plurality Elections**

<i>(a)</i> United Kingdom General Election (1997)				
	Folkestone & Hythe		Enfield Southgate	
	Votes	Share (%)	Votes	Share (%)
Conservative	20,313	39	19,137	41
Labour	12,939	25	20,570	44
Liberal Democrat	13,981	27	4,966	11

<i>(b)</i> New York Senatorial Election (1970)			
Candidate	Votes	Share (%)	
James Buckley	2,179,640	35.4	
Richard Ottinger	2,171,232	35.3	
Charles Goodell	1,434,472	23.3	

*Note:* Votes and vote shares are recorded for the three most popular candidates. The result of the New York election is taken from the election results tables at <http://www.nyhistory.net>

Labour and Liberal Democrat candidates. Nevertheless, the combined votes for them (representing 52 per cent of those who voted) would have been sufficient to dislodge Howard, had these voters coordinated behind one of the challengers. More precisely, this outcome required a fraction  $[39/52 \text{ per cent}] = 75 \text{ per cent}$  to coordinate. This magic figure is a ‘qualified majority’ of the anti-Conservative electorate. We can view, therefore, the anti-Conservative tactical coordination problem as one of qualified majority voting—a qualified majority must act together in order to achieve a commonly desired outcome. We illustrate this idea graphically in Figure 1. We write  $p$  for the proportion of anti-Conservatives who vote for Labour, and  $\bar{p}$  for the critical coordination level, so that  $\bar{p} = 0.75$  and  $p = [27/52 \text{ per cent}] = 0.52$ . The space  $[0, 1]$  divides neatly into ‘win zones’ for the Liberal Democrat, Conservative, and Labour candidates.

In Folkestone and Hythe, there was insufficient coordination to eject the incumbent Tory. In other constituencies, however, the anti-Conservative vote was concentrated on only one of the challenging candidates. In the Enfield Southgate constituency, leading Conservative politician Michael Portillo was defeated by Labour’s Stephen Twigg (Table 1(a)).

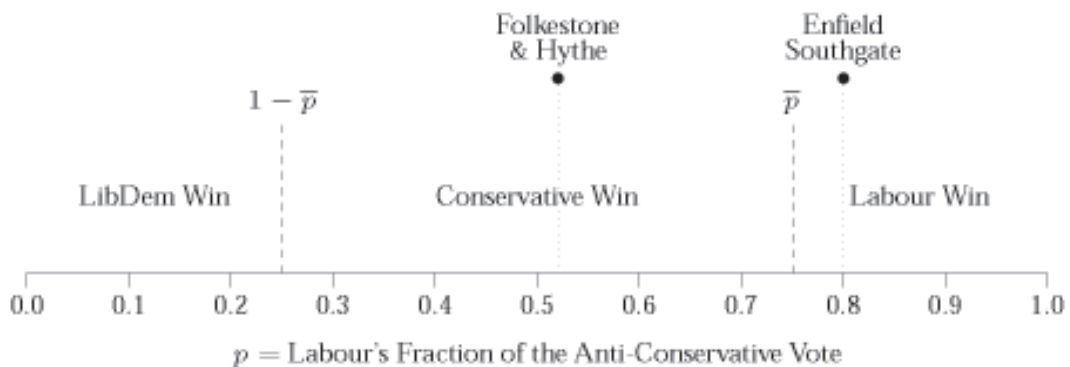
Restricting once again to the major three parties, the anti-Conservative voters needed to achieve a tactical coordination of  $\bar{p} = [41/55 \text{ per cent}] = 0.75$ —the same hurdle as Folkestone and Hythe. In this case they did so (with  $p = [44/55 \text{ per cent}] = 0.8$ ), and Portillo was ejected into the political wilderness.

In both examples, the Conservatives polled between a third and half of the votes cast for the major parties. With less than half the vote, it was at least possible for them to be defeated. Furthermore, receiving more than one-third of such votes ensured that an even split between the challengers would allow the Conservative candidates to win the seat. This situation is common, and in 1997 the Conservatives polled between one-third and a half of the major party vote in 270 English constituencies (see Figure 6). These voting districts were, therefore, a ‘battleground’ for tactical voting.

This tactical coordination problem is not unique to the United Kingdom. To a foreign observer of the United States, it might appear that all elections are a two-horse race between candidates nominated by the Democratic and Republican parties.<sup>9</sup> In New York state this is not always the case. There are a number of other smaller political parties, and at

<sup>9</sup> Such nominated candidates arise from the conduct of primary elections and these do, of course, feature more than two candidates.

**Figure 1**  
**The Tactical Coordination Problem in Plurality Voting**



times they have had a distinct influence. The 1970 New York senatorial election provides a dramatic illustration that is familiar to many political scientists.<sup>10</sup> The two familiar parties were supplemented by the Liberal and Conservative parties. In many cases, these parties nominated the Democrat and Republican candidates, yielding the usual bipartite competition.<sup>11</sup> In 1970, however, the liberals Richard L. Ottinger and Charles E. Goodell competed against the conservative James R. Buckley. Goodell was an incumbent Republican senator who had taken a liberal stance on the Vietnam War, and on this occasion received the nomination of the Liberal Party, whereas Ottinger was the Democratic nominee. The New York Conservative Party, however, rather than nominating Goodell as a ‘fusion’ candidate instead offered its support to Buckley. The outcome was a split in the liberal vote (see Table 1(b)). Liberal voters needed to coordinate their support behind either Ottinger or Goodell in order to defeat Buckley. The ‘qualified majority’ of liberal votes needed to do this was  $\bar{p} = 35.4/58.6$  per cent = 0.6. This situation is similar to that seen in our British examples, albeit with a reduced ‘hurdle’ to the coordination problem.

### III. OPTIMAL VOTING STRATEGIES

The stylized tactical voting problem described in section II is one in which (for the UK examples) the anti-Conservative voters wish to achieve, in agree-

gate, a sufficient degree of coordination. Any coordination, however, must arise from the decisions of individual voters. We must calculate, therefore, the optimal voting strategy of an instrumentally rational voter. By this we mean a voter who cares only about influencing the outcome of the election.

Of course, an individual can only exert influence when he is *pivotal*.<sup>12</sup> This happens when there is a tie for the lead (one more vote can pick the winner) or when the leading candidates are separated by a single vote (one more vote can force a tie). In all other circumstances a single vote has no effect. Our voter must consider the relative likelihood of the different possible pivotal events. We must ask what he believes about the support for each candidate.

In the context of our ‘defeat the Tory’ scenario we further simplify our analysis by supposing that the level of Conservative support (and hence the target hurdle  $\bar{p}$  for anti-Conservative coordination) is known. Restricting to voters who dislike the Tories, we write  $\pi$  for the proportion who rank the Labour candidate as their favourite. We suppose that  $\pi$  is uncertain, and model our voter’s beliefs with a probability density function  $f(\pi)$ —Figure 2 corresponds to a voter who believes that the Labour candidate is more popular.

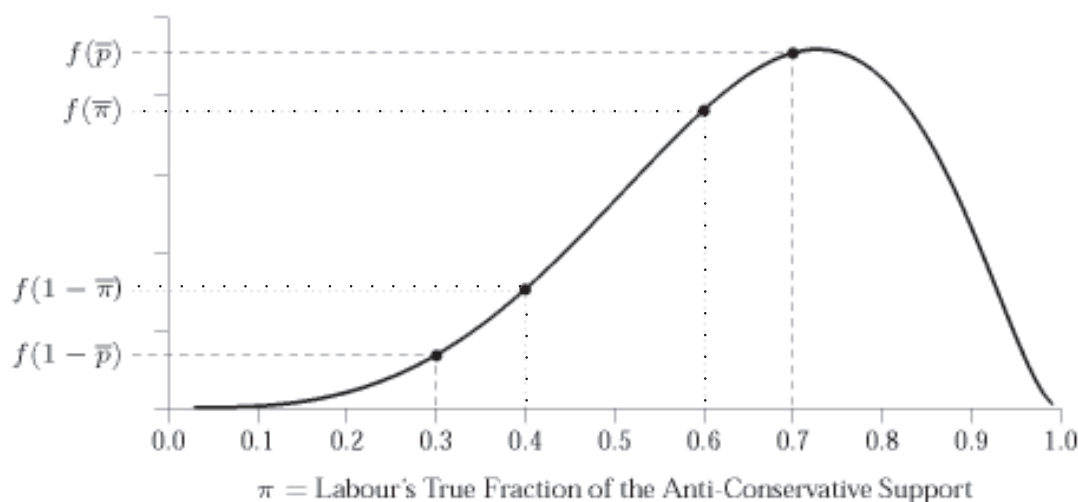
We must also specify our voter’s preferences. He gains a payoff of  $u_1 > 0$  for a Labour win,  $u_2 > 0$  for a Liberal win, and zero for a Tory win. The ratio  $u_2/$

<sup>10</sup> In a recent textbook, Morton (2001) uses this election as the canonical example of a three-horse race.

<sup>11</sup> New York State Law allows candidates to receive nominations from more than one party. In the Presidential elections of 1972, 1976, and 1984 there were coordinated nominations.

<sup>12</sup> This logic has long been familiar to political scientists. Classic treatments are offered by Riker and Ordeshook (1968), McKelvey and Ordeshook (1972), and Hoffman (1982).

**Figure 2**  
**The Relative Likelihood of a Pivotal Vote**



$u_1$  is his relative preference for the two challengers. If the voter is a true Liberal supporter, then  $u_2 > u_1$  and hence  $u_2/u_1 > 1$ .

Finally, we must specify the expected behaviour of other voters. We begin by supposing that all other voters act sincerely. This means that  $p = \pi$ —we will lift this restriction in due course.

With these ingredients in place, we may solve the individual's tactical voting problem. There are two pivotal events: either the Labour candidate will tie with the Tory ( $p = \pi = \bar{p}$ ) or the Liberal Democrat will tie with the Tory ( $p = \pi = 1 - \bar{p}$  or equivalently  $1 - \pi = \bar{p}$ ). For the beliefs illustrated in Figure 2, a Conservative–Labour tie is more likely. An extra vote for Labour will, with relative likelihood  $f(\bar{p})$ , turn a Tory win into a Labour win, resulting in a gain of  $u_1$ . Similarly, an extra Liberal vote generates a gain of  $u_2$  with relative likelihood  $f(1 - \bar{p})$ . Thus, the optimal voting strategy is to vote for Labour whenever  $u_1 f(\bar{p}) > u_2 f(1 - \bar{p})$ . Re-arranging this inequality, we obtain:

$$\frac{f(\bar{p})}{f(1 - \bar{p})} > \frac{u_2}{u_1}. \quad (1)$$

This central trade-off is as follows: an individual must balance the relative likelihood of pivotal events (the *odds*) against his relative preference for the challengers. When the Labour candidate is believed to be the stronger (so that  $f(\bar{p}) > f(1 - \bar{p})$ ), a voter who actually prefers the Liberal ( $u_2 > u_1$ ) may

nevertheless vote Labour. If he does so, then his vote is *tactical*—the relative influence of his vote outweighs his true political preferences. This is likely to be true when the odds are high and when the voter is relatively indifferent between the challengers.

This analysis is, of course, subject to a number of critiques, and we highlight two here. First, in a large electorate the *absolute* probability of a tie is extremely small.<sup>13</sup> So why vote? Any expected gain from voting will be easily outweighed by a small cost. For this reason, models based on instrumental rationality find it difficult to explain voter turnout. Possible explanations for the act of voting might then include a direct positive benefit from the act of voting, such as a feeling of democratic or party duty. Such explanations would then seem to suggest that we abandon instrumental motivations. Nevertheless, we can respond by noting that people do vote tactically, and thus appear to have instrumental motivations in mind. Furthermore, while the probability of influencing the election is small, it is not zero. In larger elections (where the influence of a single vote is tiny) the stakes are likely to be higher. Finally, once a voter decides to enter the ballot box, he may as well vote optimally, since at that point pivotal situations may be all that matter to him. Once he does so, he must balance his personal preferences against the *relative* chances (the odds, and not the *absolute* probability) of influencing the election outcome in different ways.

<sup>13</sup> In fact the absolute probability of being pivotal is inversely related to the size of the electorate.

A second critique is that our initial analysis is decision-theoretic: we assumed that other individuals voted for their true first preference, so that  $p = \pi$ . We view voting as a game, and hence an individual will anticipate strategic switching by other voters. He will expect, therefore, that the true support for a challenging candidate (reflected in  $\pi$ ) will be distorted into a different level of support at the ballot box ( $p$ ). Perhaps surprisingly, we wish to argue that such a consideration (anticipating tactical voting by others) will *reduce* an individual's incentive to vote tactically. Put simply, we claim that tactical voting may exhibit a form of *negative feedback*.

We offer the following informal argument. Suppose that the Labour candidate is the leading challenger (so that Labour's true support among anti-Conservative voters is  $\pi > 1/2$ ). The majority of voters will realize this, and hence some will tactically switch their vote away from the Liberal Democrat and towards Labour, yielding  $p > \pi$ . Similarly, when the Liberal Democrat is the leading challenger ( $\pi < 1/2$ ) we would expect  $p < \pi$ . This means that for Labour to reach the coordination target of  $p = \bar{p}$  it need not enjoy true support of  $\pi = \bar{p}$ . Instead, a smaller bias in its favour of (say  $\pi = \bar{\pi} < \bar{p}$ ) will be sufficient to do this, since the expansion in support at the ballot box (via tactical switching) will enable  $p$  to reach the target level. Similarly, a small Liberal bias will be enough for a Liberal–Conservative tie. This logic suggests that the pivotal support levels will be  $\pi = \bar{\pi}$  and  $\pi = 1 - \bar{\pi}$  where  $1/2 < \bar{\pi} < \bar{p}$ . An individual will then find it optimal to vote Labour when:

$$\frac{f(\bar{\pi})}{f(1 - \bar{\pi})} > \frac{u_2}{u_1}. \quad (2)$$

We illustrate this idea in Figure 2. Notice that (when Labour is viewed as more popular) the odds in equation (2) are smaller than those in equation (1), and a Liberal supporter faces less pressure to vote tactically. This hesitancy suggests that, in equilibrium, there will be only partial coordination of anti-Conservative voting.

We offer the following intuition. When others in the electorate are expected to switch their votes, then it takes a smaller bias in favour of one challenger to catch the Tories. The electoral situation is, in some sense, more precarious, and hence an individual

voter is more cautious in his decisions. In particular, the voter must be careful not to follow the wrong tactical voting bandwagon. Although he may perceive a Labour lead, he recognizes that others may perceive a Liberal Democrat lead, and hence he is more hesitant to switch his own vote.

This argument leans upon a number of suppositions. For instance, a voter who believes that the Labour candidate is stronger (as in Figure 2) must envisage the possibility that  $\pi = 1 - \bar{\pi} < 1/2$ . He recognizes that the Liberal may be the true leading challenger and, importantly, that other voters may believe that  $\pi < 1/2$ . This means that he is uncertain of the game that he is playing, and of the game that other voters think that they are playing. There is no *common knowledge* of the constituency situation—it is this feature that makes voting a global game.

#### IV. TACTICAL VOTING EQUILIBRIA IN PLURALITY ELECTIONS

We have suggested that a voting equilibrium will involve only partial tactical vote switching. This implies variation in voter behaviour, with some voters choosing to switch and some casting a sincere vote. An optimal voting decision involves a trade-off between a voter's preferences and the odds of influencing the election outcome. This means that different individuals will act in different ways for two reasons: first, they may have different preferences and, second, they may perceive different odds of influencing the election outcome. We wish to focus on the second of these factors, and must ask why voters may hold different beliefs.

The extensive media coverage of modern elections suggests that voters are likely to share similar expectations. Such coverage, however, is often limited to *national* party support. In the United Kingdom, votes are aggregated at the *constituency* level to determine Members of Parliament, and typically the largest party in parliament then goes on to form an executive. A voter must consider, therefore, the support of the different candidates within her own constituency, rather than assessing only the national situation. National media coverage gives common information to people across the country, but rarely gives widely observed reports on indi-

**Box 1: A Formal Model of Voter Communities**

Formally, our constituency is a continuum of communities  $i \in [0, 1]$ , similar in spirit to the ‘islands’ model employed by Amato, Morris, and Shin, in their article in this issue. Restricting to anti-Conservative voters, Labour’s support  $\pi_i$  in community  $i$  is drawn independently from the density  $g(\pi_i | \pi)$ , and

$$\pi = \int_0^1 \pi_i di.$$

An individual does not know the constituency situation  $\pi$ , but has prior beliefs represented by the density  $f(\pi)$ . Observing  $\pi_i$ , she Bayesian updates to obtain a posterior:

$$f(\pi | \pi_i) = \frac{f(\pi)g(\pi_i | \pi)}{\int_0^1 g(\pi_i | \pi')f(\pi')d\pi'} \Rightarrow \frac{f(\pi | \pi_i)}{f(1 - \pi | \pi_i)} = \frac{f(\pi)}{f(1 - \pi)} \frac{g(\pi_i | \pi)}{g(\pi_i | 1 - \pi)}. \quad (3)$$

Thus a voter’s posterior odds of  $\pi$  versus  $1 - \pi$ , conditional on her community situation  $\pi_i$ , are equal to the prior odds multiplied by the likelihood ratio of observing  $\pi_i$  in each circumstance. We (implicitly) use three assumptions in our analysis.

Assumption 1. The prior is symmetric, so that  $f(\pi) = f(1 - \pi)$ .

Assumption 2. For  $\pi > 1/2$  the likelihood ratio  $g(\pi_i | \pi)/g(\pi_i | 1 - \pi)$  is increasing in  $\pi_i$ .

Assumption 3. For  $\pi_i > 1/2$  and  $\pi > 1/2$  the likelihood ratio is decreasing in  $\pi$ .

Assumptions 2 and 3 are technical monotone likelihood ratio properties that are satisfied by a wide range of specifications, including the nested probit model described in Box 2. Assumption 1, however, rules out any prior bias in a voter’s beliefs. Given these assumptions, and equation (3), a Liberal supporter would find it optimal to switch her vote to Labour whenever:

$$\frac{g(\pi_i | \bar{\pi})}{g(\pi_i | 1 - \bar{\pi})} > U.$$

The prior odds are eliminated by Assumption 1. The likelihood ratio is increasing in  $\pi_i$ , by Assumption 2. Thus there is a threshold  $\bar{\pi}$  beyond which a voter will find it optimal to vote for the Labour candidate, even though she prefers the Liberal Democrat. Using Assumption 3, this threshold decreases as the required true support  $\bar{\pi}$  for a win by Labour increases.

vidual constituencies.<sup>14</sup> There is scope for individual voters to hold widely different beliefs about their constituency situation.

Variation in voter beliefs may arise from differences in the information sources available to them. For instance, day-to-day interaction with other voters throughout a constituency will give an individual a picture of the support enjoyed by each candidate. This picture is unlikely to be complete, however. We might imagine that individuals tend to interact with others like them, perhaps within conceptual communities. Social interaction, therefore, enables voters

to work out a candidate’s support within a community, but not across an entire constituency. Thus, as we move through different communities we might expect voters’ beliefs to vary quite markedly.

The idea of community variation is not the only way to generate differences in voter opinions, but nevertheless is useful. We formalize it in the following way.<sup>15</sup> We divide our constituency into communities, indexed by  $i$ . Among anti-Conservative voters, a proportion  $\pi_i$  prefer Labour in community  $i$ . The constituency-wide support  $\pi$  is then the average across all communities. For simplicity, we suppose

<sup>14</sup> During the UK General Election of 1997, a large number of opinion polls were conducted throughout the campaign. Only 29 of these were conducted at a constituency level, covering 26 out of a total of 659 constituencies (Evans *et al.*, 1998). Thus opinion polls failed to give voters an accurate picture of their constituency situation.

<sup>15</sup> We relegate our formal analysis to Box 1, and offer a parametric illustration in Box 2.

that the intensity of voter preferences does not vary.<sup>16</sup> A voter enjoys a payoff of  $U$  for her first preference, 1 for her second preference, and zero for a Tory win. The different payoff configurations and their proportions in community  $i$  are therefore:

	$\pi_i$	$1 - \pi_i$
Labour ( $u_1$ )	$U$	1
Liberal Democrat ( $u_2$ )	1	$U$
Conservative	0	0

A voter in community  $i$  observes  $\pi_i$  but not  $\pi$ , and uses this observation to form her beliefs about the overall support  $\pi$  for the two different challengers. Some voters will believe that the Labour candidate is the leading challenger ( $\pi_i > 1/2$ ), while others will think that the Liberal Democrat is stronger ( $\pi_i < 1/2$ ). Fixing  $\pi$ , when  $\pi_i$  varies widely between different constituencies, then the voters will tend to hold different beliefs. Similarly, when communities are homogeneous (for instance,  $\pi_i = \pi$  for all  $i$ ) then voters will all view the constituency situation in the same way. Interestingly, the presence of heterogeneous communities (i.e. variation in  $\pi_i$ ) helps explain why voters may tend to overestimate the likelihood of success of their preferred candidate. Suppose that the Liberal Democrat is less popular than the Labour candidate ( $\pi > 1/2$ ). A voter who prefers the Liberal Democrat is more likely to be found in a community where the Liberals are more popular than in the constituency as a whole ( $\pi_i < \pi$ ). This means that such voters are likely to receive information that is biased in favour of their true first preference. Nevertheless, from their own perspective, this information is unbiased. Thus the voter expectations (see section I) described by Norris *et al.* (1999) are necessarily inflated. In other words, there is likely to be correlation between a voter's information sources and her preferences, and thus correlation between her beliefs and her true preferences.

We wish to relate the information available to a voter (summarized by  $\pi_i$ ) to the decision that she takes. Consider a voter who actually prefers the Liberal candidate. We suppose that she follows a

*threshold rule* in making her voting decision. We will later show that such a rule corresponds to an optimal voting strategy. A threshold rule works as follows. Our voter acts sincerely (voting for the Liberal Democrat) unless she observes support for the Labour candidate  $\pi_i$  that exceeds some threshold  $\bar{\pi}_i \geq 1/2$ , in which case she votes tactically. Similarly, a Labour supporter will switch to the Liberals when  $\pi_i < 1 - \bar{\pi}_i$ . Such a threshold rule is illustrated in Figure 3(a) for  $\bar{\pi}_i = 0.65$ . The threshold  $\bar{\pi}_i$  indexes the degree of tactical voting. For  $\bar{\pi}_i = 1$ , an individual will always vote sincerely. In contrast, for  $\bar{\pi}_i = 1/2$ , an individual will vote for the most popular challenger. Thus voters act sincerely unless they observe a severe bias in favour of one of the challengers.

For a given threshold rule, we can aggregate voting behaviour across communities, and arrive at the relationship between true constituency support  $\pi$  and vote share  $p$  illustrated in Figure 3(b). Notice that an increase in  $\pi$  has two effects. First, it increases the overall level of Labour support. Second, it increases the number of constituencies in which voters switch tactically to Labour and reduces the number of constituencies in which voters switch tactically to the Liberal Democrats. Using Figure 3(b), we can calculate the support  $\bar{\pi}$  needed for Labour's vote share  $p$  to hit the required target  $\bar{p}$ . A small lead for Labour is expanded via tactical voting, and hence  $\bar{\pi} < \bar{p}$ . Furthermore, as  $\bar{\pi}_i$  decreases—so that individuals are more willing to vote tactically—the required support level decreases too—see Figure 3(c). Thus we have (Figure 3(d)) a positive relationship between  $\pi$  and  $\bar{\pi}_i$ . Notice that for  $\bar{\pi}_i = 1$  we need  $\pi = \bar{p}$  for a Tory defeat.

This analysis tells us the critical level of support  $\bar{\pi}$  that a challenger needs in order to defeat the disliked Tory, given that a threshold voting rule is in operation. Of course, it does not tell us whether the threshold rule itself is optimal. We turn, therefore, to this second issue. Consider a voter who prefers the Liberal candidate (so that  $u_2 = U$  and  $u_1 = 1$ ) and observes a fraction  $\pi_i > 1/2$  support for the Labour candidate in her community. Based on  $\pi_i$ , she may form beliefs  $f(\pi | \pi_i)$  about Labour's support. Our

<sup>16</sup> This means that any variation in tactical voting behaviour is entirely due to differences in the beliefs held by voters. Of course, a more realistic model would allow for variation in the intensity of preferences as well. Myatt (2002a,b) allows for both factors, and generates similar results.

**Box 2: A Nested Probit Model of Voter Communities**

We illustrate our ‘island communities’ model with a simple nested probit specification. We suppose that anti-Conservative voter  $j$  in community  $i$  is characterized by  $x_{ij}$ , preferring Labour if and only if  $x_{ij} > 0$ . We generate  $x_{ij}$  by nesting normal distributions:

$$x_i \sim N(\theta, \sigma^2) \text{ and } x_{ij} | x_i \sim N(x_i, 1 - \sigma^2).$$

Thus  $\theta$  represents the median voter in the constituency. Individuals differ from this median via a community effect ( $x_i - \theta$ ) and an individual effect ( $x_{ij} - x_i$ ), so that  $x_i$  is the median voter in community  $i$ . Thus  $\sigma^2$  represents variance *between* communities, whereas  $1 - \sigma^2$  is the variance *within* communities. Combining these effects, we have  $x_{ij} \sim N(\theta, 1)$ . When  $\sigma^2$  is small, most variation is within the communities, and hence  $x_i$  will be close to  $\theta$ . This means that an observation of support within a community gives a relatively good picture of the constituency situation. Similarly, when  $\sigma^2$  is large, communities are relatively different, and give an imperfect picture of the community situation. We may translate the median voters  $\theta$  and  $x_i$  into constituency and community support respectively. To do this, we write  $\Phi(\cdot)$  and  $\phi(\cdot)$  for the cumulative distribution and density functions of the standard normal distribution. Hence:

$$\pi = \Phi(\theta) \text{ and } \pi_i = \Phi\left(\frac{x_i}{\sqrt{1 - \sigma^2}}\right).$$

We can also invert these relationships to obtain the identity of the median voter in terms of support levels, so that

$$\theta = \Phi^{-1}(\pi) \text{ and } x_i = \sqrt{1 - \sigma^2} \Phi^{-1}(\pi_i).$$

We may use this specification to consider a voter’s beliefs over  $\pi$  given the community support  $\pi_i$ . Adopting a uniform prior ( $f(\pi) = 1 \forall \pi$ ) for simplicity, the posterior beliefs of a voter in community  $i$  satisfy:

$$f(\pi | \pi_i) \propto \phi\left(\frac{\sqrt{1 - \sigma^2} \Phi^{-1}(\pi_i) - \Phi^{-1}(\pi)}{\sigma}\right).$$

For  $\sigma^2 = 0.25$ , this is the formulation used to generate the density in Figure 2. We can use this formulation to consider the relative influence of a vote for Labour versus the Liberal Democrats. Suppose that the critical support levels to generate Labour and Liberal Democrat wins are  $\bar{\pi}$  and  $1 - \bar{\pi}$  respectively. Then simple algebraic manipulations (see the Appendix) reveal that:

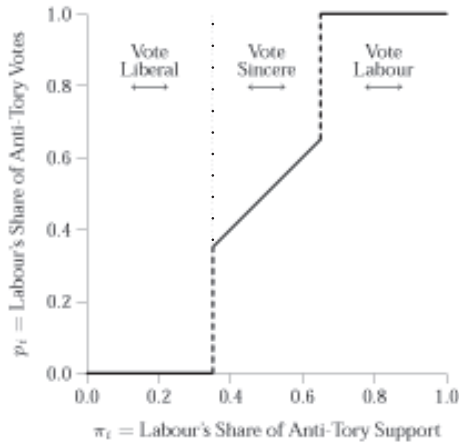
$$\frac{f(\bar{\pi} | \pi_i)}{f(1 - \bar{\pi} | \pi_i)} = \exp\left(\frac{2\sqrt{1 - \sigma^2} \Phi^{-1}(\pi_i) \Phi^{-1}(\bar{\pi})}{\sigma^2}\right).$$

For  $\bar{\pi} > 0$ , this expression is increasing in  $\pi_i$ . This means, quite simply, that there is a stronger incentive to vote for Labour in communities in which Labour’s support is stronger. Furthermore, there will be a critical support level  $\bar{\pi}_i$  for which it is optimal for a Liberal Democrat supporter to switch her vote to Labour. More algebra (see again the Appendix) reveals that:

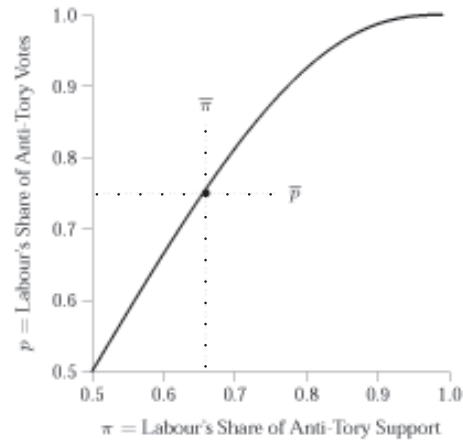
$$\bar{\pi}_i = \Phi\left(\frac{\sigma^2 \log U}{2\sqrt{1 - \sigma^2} \Phi^{-1}(\bar{\pi})}\right).$$

This relationship is illustrated in Figure 4(b).

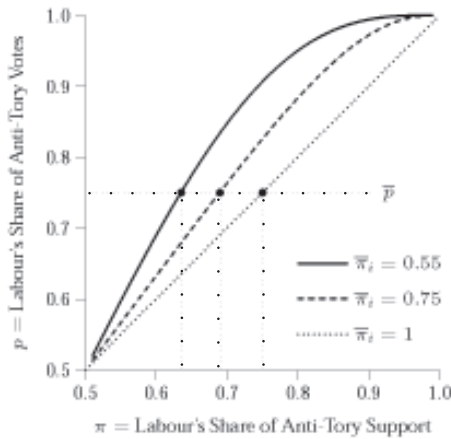
**Figure 3**  
**The Translation of Candidate Support to Votes at the Ballot Box**



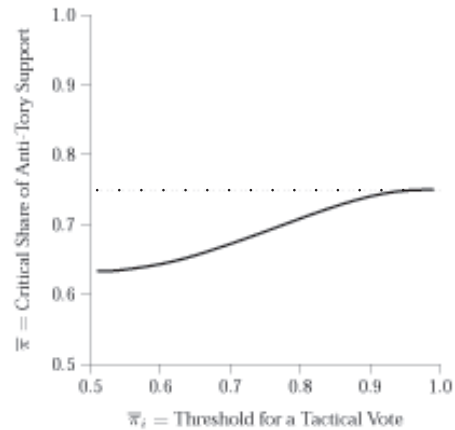
(a) From support to votes in the community



(b) From support to votes in the constituency



(c) Calculating  $\bar{\pi}$  from different  $\bar{\pi}_i$



(d) Relationship between  $\bar{\pi}_i$  and  $\bar{\pi}$  for  $\bar{p} = 0.75$

*Notes:* Figure 3(a) shows the relationship in community  $i$  between true Labour support  $\pi_i$  and vote share  $p_i$ , when voters use a threshold rule  $\bar{\pi}_i = 0.65$ . For  $0.35 < \pi_i < 0.65$ , voters observe a relatively even anti-Tory split between Labour and the Liberal Democrats, and so  $p_i = \pi_i$ . For  $\pi_i > 0.65$ , all Liberal Democrats are sufficiently swayed by Labour's popularity to vote tactically, and hence  $p_i = 1$ . Similarly,  $p_i = 0$  for  $\pi_i < 0.35$ . Figure 3(b) aggregates this relationship across the constituency. For this example we have used the nested probit specification from Box 2 with  $\sigma^2 = 0.25$ , and a threshold rule of  $\bar{\pi} = 0.65$ . This allows us to calculate the true support  $\bar{\pi}$  for Labour among anti-Conservatives needed for Labour's vote share to hit  $\bar{p}$  and hence defeat the Tory. Increasing  $\bar{\pi}_i$  (so that voters are less willing to vote strategically) increases  $\bar{\pi}$ , as illustrated in Figures 3(c) and 3(d).

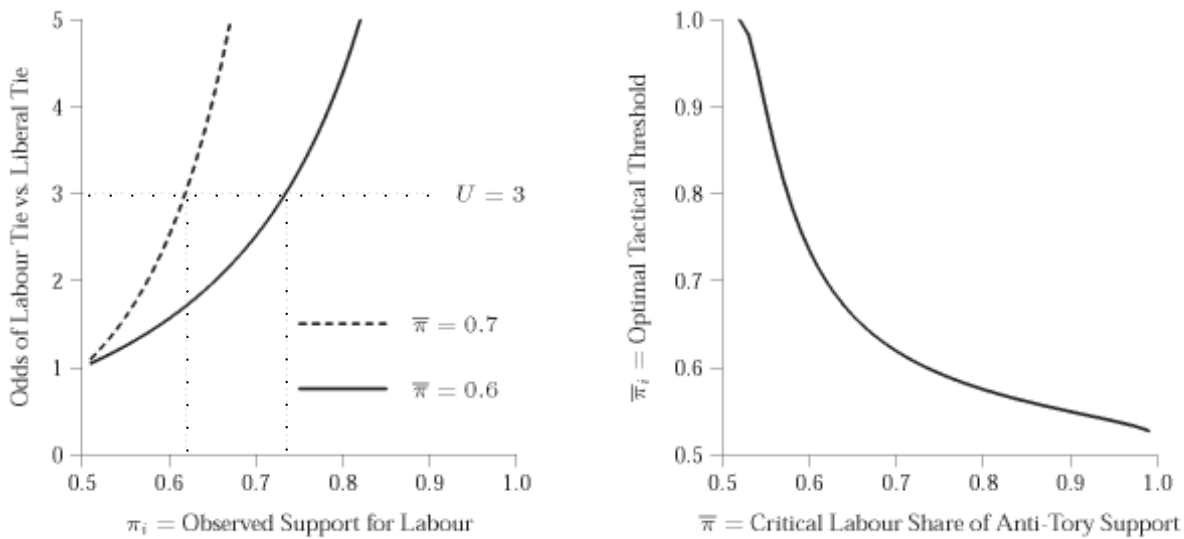
earlier analysis of the optimal voting problem tells us that she should switch her vote to Labour whenever:

$$\frac{f(\bar{\pi} | \pi_i)}{f(1 - \bar{\pi} | \pi_i)} > U,$$

in other words, when the odds of influencing the election in Labour's favour (evaluated following the informational input from her community situation  $\pi_i$ )

exceed her relative preference. For a wide variety of specifications (see Boxes 1 and 2) these odds are increasing in  $\pi_i$ . So, when Labour is perceived to be strong it is relatively more likely that Labour will tie with the Conservatives—see Figure 4(a) for an illustration. This means that our voter will find it optimal to vote tactically whenever  $\pi_i > \bar{\pi}_i$ . This is precisely the threshold rule that we had originally

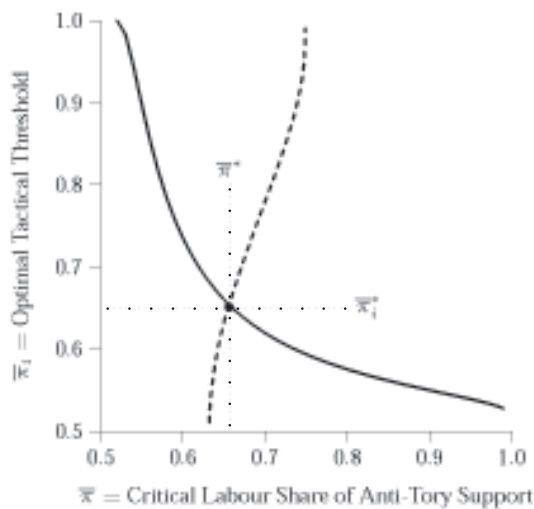
**Figure 4**  
**Optimal Tactical Voting Thresholds**



(a) Relative preference versus posterior odds

(b) Effect of  $\bar{\pi}$  on threshold  $\bar{\pi}_i$

**Figure 5**  
**Tactical Voting Equilibrium**



envisaged. Furthermore, under reasonable assumptions the odds will also be increasing in  $\bar{\pi}$ . Inspecting Figure 4(a) once more, this means that when  $\bar{\pi}$  is higher, a voter will find it optimal to use a lower threshold  $\bar{\pi}_i$ . Thus we obtain the relationship illustrated in Figure 4(b).

We now have two relationships that may be combined to characterize a voting equilibrium. Given that Labour needs to enjoy a share  $\bar{\pi}$  of anti-Tory support in order to win, Figure 4(b) gives us the optimal threshold rule that will be used by voters.

Given that voters use such a rule with threshold  $\bar{\pi}_i$ , Figure 3(d) tells us in turn the support  $\bar{\pi}$  that Labour needs to enjoy across the community. For an equilibrium, therefore, we need a pair  $\bar{\pi}$  and  $\bar{\pi}_i$  that satisfies both of these relationships. In Figure 5, the intersection of the two curves generates a tactical voting equilibrium.

We offer a number of observations. First, the equilibrium is unique.<sup>17</sup> Second, the ‘threshold’ strategies are easily interpreted—a voter is swayed only when her second-choice candidate is sufficiently

<sup>17</sup> More precisely, it is unique within the class of equilibria that we consider.

popular. Third, the equilibrium involves only partial strategic voting, since  $1/2 < \bar{\pi}_i^* < 1$ . This means that voters in communities where the anti-Conservative support is evenly split ( $\pi_i \approx 1/2$ ) will vote sincerely. Fourth, voters are less willing to vote tactically than they would be if they expected all others to vote truthfully. Fifth, our model offers potential comparative statics. Varying parameters such as the intensity of preferences  $U$  and the required ‘hurdle’ for anti-Conservative coordination  $\bar{p}$  will shift the curves in Figure 5 and hence the equilibrium solution.

We do not perform a formal investigation of the comparative statics here, but we can confirm that the model generates results that reflect those generated from our other work (Myatt, 2002a; Myatt and Fisher, 2002). We find that tactical voting increases (in the sense of a lower equilibrium threshold  $\bar{\pi}_i^*$ ) when voters are more indifferent between the challengers (lower  $U$ ), when the coordination hurdle is higher (higher  $\bar{p}$ ), and when the information available to voters is less variable (which corresponds to a lower value for  $\sigma^2$  in the nested probit model of Box 2). Interestingly, we can also use the model to *invert* election results. Given the parameters of the model, the solution maps true support for challenging candidates (reflected in  $\pi$ ) to vote share at the ballot box (or  $p$ ). Given the outcome to an election  $p$ , we can invert this relationship to find the true support  $\pi$ . This gives us insight into what the election result would have been in the absence of any tactical voting. We turn, therefore, to consider a particular calibration of our model, enabling us to comment on the policy significance of tactical voting.

## V. THE IMPACT OF TACTICAL VOTING

Our ‘defeat the Tory’ model is a stylized representation of the problem faced by anti-Conservative voters in recent UK General Elections. In 1997 the desire to overturn the incumbent administration was at its highest. As we have already noted, the constituencies of England were three-horse races between the main political parties. In Figure 6, we use

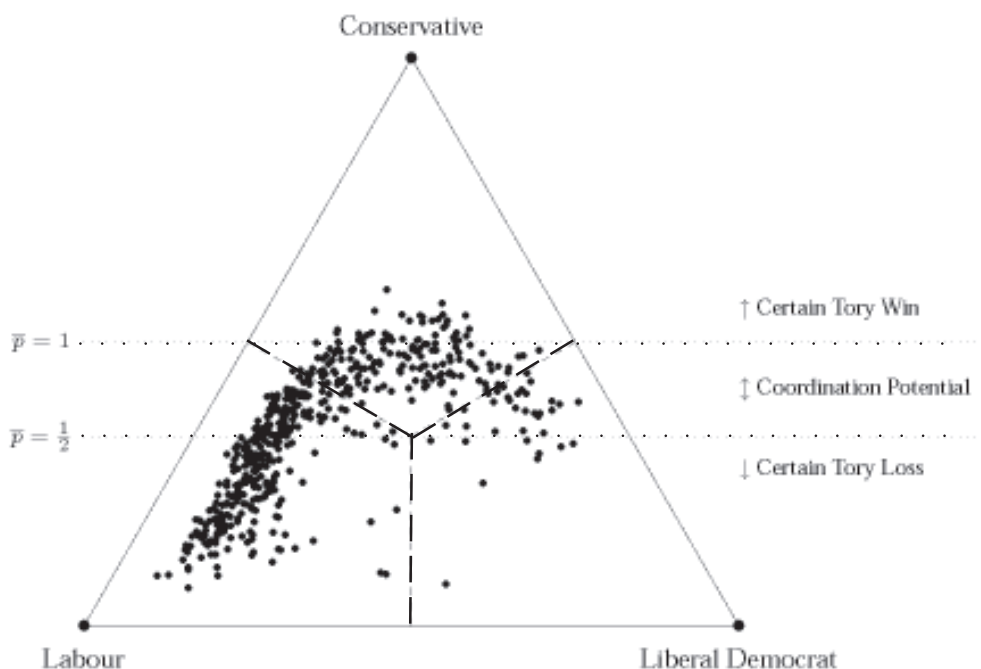
a barycentric plot to display the outcomes from 527 English constituencies. The game applies when the Conservative candidate enjoys more than one-third but less than half of the votes cast for the three main parties. Inspecting Figure 6, we can see that there were 270 such constituencies in 1997, and these were the ‘battleground’ for tactical voting. Of the remaining seats, the vast majority were very safe Labour seats (from the bottom left of the diagram) and hence the 270 battleground constituencies are the focus of our interest.

We apply our model to these constituencies by using the following calibration procedure. First, for each constituency, we calculate the coordination hurdle  $\bar{p}$  needed to defeat the Tory. Second, we fix the relative preferences of anti-Conservative voters. For this exercise, we chose  $U = 2$ , so that a voter rates his first choice twice as highly as his second, at least once we have normalized the payoff of a Tory win to zero. Third, given a split  $\pi$  in anti-Conservative support, we must specify the distribution of support across different communities within a constituency. We do this using the nested probit specification of Box 2, and pick a parameter of  $\sigma^2 = 0.25$ . This says that of variation in voter characteristics across a constituency, 25 per cent occurs between communities and 75 per cent occurs within communities. With these parameters in hand, we can compute the equilibrium tactical voting threshold for each community.<sup>18</sup>

The results of our calibration exercise are shown in Table 2. Previous studies have assessed the change between elections to determine which seats that changed hands did so as a result of tactical voting. This is the first serious attempt to estimate the *total* number of seats won or lost by tactical voting at a single election. It is, however, comforting to note that the results are within the range expected from the Nuffield studies if we assume that Labour and the Liberal Democrats have accumulated seats as a result of tactical voting at each election from 1987 to 1997 (Curtice and Steed, 1988, 1992, 1997). Our analysis (Table 2(a)) shows that tactical voting cost the Conservatives 46 seats, of which 35 were taken by Labour and 11 by the Liberal Democrats. For

<sup>18</sup> The parameters chosen are purely illustrative. Nevertheless, they are in fact consistent with an application of this and related models to survey data. See, for instance, Myatt (2002a) and Myatt and Fisher (2002).

**Figure 6**  
**Results from the British Election of 1997**



*Notes and Source:* This figure is reproduced from Myatt (2002a). The relative vote shares for the three major parties for 527 English constituencies are plotted using a *barycentric* or *simplex* plot. The three corners of the simplex represent 100 per cent vote shares for the labelled party. A bullet point ‘•’ indicates a constituency. Its location on the graph is a weighted average of the three extreme points, with weights corresponding to the relative vote shares of the major parties. The sides of the simplex (solid lines) represent points where only two parties receive votes. The dashed lines represent points where two parties tie for the lead, and hence separate the ‘win zones’ for each party. The dotted lines delineate the constituencies in which the Conservative Party polled between one-third and a half of the votes for major parties. These are the 270 constituencies in which coordination of anti-Conservative voters is required to avoid a Tory win.

Labour in 1997, 35 extra seats could be thought of as icing on the cake. However, a similar analysis of 1992 shows that John Major’s government would not have been in such a precarious position with a narrow majority had the Conservatives been able to avoid the loss of seats through tactical voting. While it is the practice of a small minority in special strategic circumstances, tactical voting is decisive in a substantial number of constituencies and can therefore have an important political impact.

Returning to the subject at the start of this paper, it remains to ask how we rate the relative importance of Duverger’s mechanical and psychological effects that are said to produce a two-party system at the national level. Table 2(b) decomposes these

effects. The first four columns give the actual votes and seats, and the predicted votes and seats under sincere voting, both expressed as a percentage for the 270 constituencies under consideration. The total effect of the electoral system is the difference between what would have happened in a perfect PR electoral system, and the actual result in terms of seats won. Since we expect electors to vote sincerely in a PR system, the sincere votes (third column) tell us the distribution of seats we should expect to find under PR. The total system effect in the final column is therefore the difference between the sincere votes and actual seats. It is clear that the FPTP electoral system has indeed benefited the leading two parties over the third-placed party, as Duverger (1954) described. Under a PR system the Liberal Democrats would have gained 24 per cent

**Table 2**  
**Results from the Calibration Exercise**

(a) Seat totals with and without tactical voting (England 1997)

	Actual	Predicted sincere	Change
Conservative	127	173	-46
Labour	118	83	+35
Liberal Democrat	25	14	+11

(b) Separation of the psychological and mechanical effects (England 1997)

	Actual (%)		Predicted sincere (%)		Electoral system effect (%)		
	votes	seats	votes	seats	mechanical	psychological	total
Conservative	41	47	41	64	+23	-17	+6
Labour	37	44	35	31	-4	+13	+9
Liberal Democrat	22	9	24	5	-19	+4	-15

*Notes:* Table 2(a) records the actual and inverted outcomes in the 270 battleground seats. The first column ('Actual') records the number of seats actually won by each party. For each seat we calculate  $\bar{p}$  and the actual split  $p$  in the anti-Conservative vote. Using our model, we then calculate the notional true split in the anti-Conservative vote  $\pi$ . Translate this parameter into vote shares once more, we re-calculate the election results to obtain the second column ('Predicted sincere'). The third column ('Change') records the difference between the two, and hence the number of seats affected by the 'psychological effect' of tactical voting. Table 2(b) is calculated in a similar way. The first and second columns give the vote and seat shares of the parties. The third and fourth columns give these shares after the computation of our 'sincere' outcomes. The remaining three columns decompose the Duvergian psychological and mechanical effects, as described in the main text.

of the seats considered here, when in fact they won only 9 per cent.

So how much of this discrepancy is due to tactical voting? The simple answer is none. Not only is all of the damage to the Liberal Democrats due to the mechanical translation of votes to seats, but tactical voting actually offsets some of this and improves the performance of the third-placed party. The total effect of the electoral system can be decomposed into the mechanical and psychological effects that Duverger described. The mechanical effect is the difference between the distribution of votes and seats assuming everyone votes sincerely (i.e. the difference between the third and fourth columns of Table 2(b)). Since the Conservatives were the largest party in the seats we consider, they come first in a constituency more often and, importantly, far more often than merited by the share of the vote.

So the mechanical effect concentrates power in Tory hands at the expense of both Labour and the Liberal Democrats.

Tactical voting—the psychological effect—partially reverses the mechanical effect for the Liberal Democrats and overcompensates the Labour party. The psychological effect is the difference between the predicted FPTP distribution of seats under sincere voting and the actual distribution of seats (i.e. columns two and four of Table 2(b)). Naturally, tactical voting distributes anti-Conservative votes more efficiently and therefore can only damage the Tories and benefit Labour and the Liberal Democrats. Since Labour was more often better placed to beat the Conservatives, Labour benefited more from tactical voting than the Liberal Democrats. Duverger's psychological effect, while clearly creating a tendency towards a two-party system at the

local level, does not do so at the national level. Indeed, tactical voting can help sustain three-party competition.

It appears, then, that Duverger mistook the ‘wasted vote’ logic of tactical voting at the local (or constituency) level to imply a detrimental effect to the third party nationally. But some may object that in addition to local tactical voting there may still be a kind of national psychological effect whereby voters respond to the weakness of the third party nationally by switching to another. This may be because the third party has little chance of forming a government (and despite substantial coalition potential). But, there is no clear evidence for such a phenomenon and since it does not make sense it should not be expected.<sup>19</sup>

Although the analysis here is based on a limited number of constituencies, the character of the conclusions, if not the magnitude of the effects, apply to the UK as a whole. The constituencies in this study are those where the potential for tactical voting to have an effect (both on the distribution of votes and seats) was the greatest. The frequency of tactical voting elsewhere will have been very much lower and correspondingly exceedingly unlikely to have had much impact. None the less it is not ideal

to assume that tactical voting is precisely zero outside the constituencies we have considered, and this explains why we have not translated the results here to their UK-wide equivalents.

While the total effect of the electoral system in 1997 is exactly as predicted by Duverger’s law (i.e. to the disadvantage of the third party) it is not the product of mechanical and psychological effects acting in the same direction. Rather it is an accident that the mechanical and psychological effects, which act in different directions, balance each other out in the way they do. Clearly the strongest of the effects is the mechanical one, and its strength depends on the distribution of the votes for each party across the constituencies. The psychological effect is also conditional on this distribution because it helps define the strategic situation in the constituencies. The fact that Labour and the Liberal Democrats are closer to each other than either is to the Conservatives explains why tactical coordination is overwhelmingly to the detriment of the Tories. But this feature of party preference structure is itself contingent. In the early 1980s the Liberals were officially half-way between Labour and the Conservatives, and the pattern of party policy proximity may change again in the future.

## APPENDIX: MATHEMATICAL PROOFS

### A Nested Probit Model of Voter Communities

In Box 2 we described a nested probit model of voter communities, and here we explain the derivation of voter beliefs. First, we derive the support for Labour in community  $i$ . Conditional on  $x_p$ , we have  $x_{ij} \sim N(x_p, 1 - \sigma^2)$  and hence:

$$\pi_i = \Pr[x_{ij} > 0 | x_i] = \Pr\left[\frac{x_{ij} - x_i}{\sqrt{1 - \sigma^2}} > -\frac{x_i}{\sqrt{1 - \sigma^2}}\right] = 1 - \Phi\left(-\frac{x_i}{\sqrt{1 - \sigma^2}}\right) = \Phi\left(\frac{x_i}{\sqrt{1 - \sigma^2}}\right).$$

A similar procedure ensures that  $\pi = \Phi(\theta)$ . Next, we need to consider the distribution of  $\pi_i$  conditional on  $\pi$ . Writing  $G(\pi_i | \pi)$  for the conditional distribution function and  $g(\pi_i | \pi)$  for the density:

<sup>19</sup> It would never be consistent to desert the Liberal Democrats if they were better placed to beat the Conservative candidate locally. In places where the Liberal Democrats come last the national and local strategic considerations are the same. The remainder are Labour–Liberal Democrat dominated constituencies. But even here, a Liberal Democrat supporter would not vote for a Labour candidate when the alternative would be a Liberal Democrat MP.

$$\begin{aligned}
G(\pi_i | \pi) &= \Pr \left[ \Phi \left( \frac{x_i}{\sqrt{1-\sigma^2}} \right) \leq \pi_i | \pi \right] \\
&= \Pr \left[ \frac{x_i - \Phi^{-1}\pi}{\sigma} \leq \frac{\sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(\pi)}{\sigma} | \pi \right] \\
&= \Phi \left( \frac{\sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(\pi)}{\sigma} \right) \\
\Rightarrow g(\pi_i | \pi) &= \frac{\sqrt{1-\sigma^2}}{\sigma\phi(\Phi^{-1}(\pi_i))} \phi \left( \frac{\sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(\pi)}{\sigma} \right).
\end{aligned}$$

We now evaluate the conditional beliefs of a voter. Bayesian updating from a prior of  $f(\pi)$ :

$$f(\bar{\pi} | \pi_i) \propto f(\pi)g(\pi_i | \pi) \propto f(\pi) \phi \left( \frac{\sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(\pi)}{\sigma} \right).$$

Evaluating the posterior likelihood of  $\pi$  versus  $1 - \pi$ , and imposing a symmetric prior ( $f(\pi) = f(1 - \pi)$ ) simplifies this further:

$$\begin{aligned}
\frac{f(\bar{\pi} | \pi_i)}{f(1 - \bar{\pi} | \pi_i)} &= \exp \left( - \frac{\left[ \sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(\pi) \right]^2 - \left[ \sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(1-\pi) \right]^2}{2\sigma^2} \right) \\
&= \exp \left( - \frac{\left[ \sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) - \Phi^{-1}(\pi) \right]^2 - \left[ \sqrt{1-\sigma^2}\Phi^{-1}(\pi_i) + \Phi^{-1}(\pi) \right]^2}{2\sigma^2} \right) \\
&= \exp \left( \frac{2\sqrt{1-\sigma^2}\Phi^{-1}(\pi_i)\Phi^{-1}(\pi)}{\sigma^2} \right).
\end{aligned}$$

The first step takes the ratio of normal density functions. The second uses the symmetry of the normal to yield  $\Phi^{-1}(1 - \pi) = -\Phi^{-1}(\pi)$ , following which the exponent simplifies nicely. Suppose now that the critical level of support necessary to generate a Labour–Conservative tie is  $\bar{\pi}$ , and that to generate a Liberal–Conservative tie is  $1 - \bar{\pi}$ . A Liberal Democrat supporter will be just indifferent between the two challenging candidates when:

$$\begin{aligned}
\frac{f(\bar{\pi} | \pi_i)}{f(1 - \bar{\pi} | \pi_i)} &= \frac{U}{1} \Leftrightarrow \frac{2\sqrt{1-\sigma^2}\Phi^{-1}(\pi_i)\Phi^{-1}(\bar{\pi})}{\sigma^2} = \log U \\
&\Leftrightarrow \Phi^{-1}(\pi_i) = \frac{\sigma^2 \log U}{2\sqrt{1-\sigma^2}\Phi^{-1}(\bar{\pi})} \\
&\Leftrightarrow \pi_i = \Phi \left( \frac{\sigma^2 \log U}{2\sqrt{1-\sigma^2}\Phi^{-1}(\bar{\pi})} \right).
\end{aligned}$$

## REFERENCES

- Amorim Neto, O., and Cox, G. W. (1997), 'Electoral Institutions, Cleavage Structures, and the Number of Parties', *American Journal of Political Science*, **41**, 149–74.
- Carlsson, H., and Van Damme, E. (1993), 'Global Games and Equilibrium Selection', *Econometrica*, **61**, 989–1018.
- Cox, G. W. (1984), 'Strategic Electoral Choice in Multi-member Districts: Approval Voting in Practice?', *American Journal of Political Science*, **28**, 722–38.
- (1994), 'Strategic Voting Equilibria under the Single Nontransferable Vote', *American Political Science Review*, **88**(3), 608–21.
- (1997), *Making Votes Count*, Cambridge, Cambridge University Press.
- Curtice, J., and Steed, M. (1988), 'Appendix 2: Analysis', in D. Butler and D. Kavanagh (eds), *The British General Election of 1987*, Basingstoke, Macmillan.
- — (1992), 'Appendix 2: The Results Analysed', in D. Butler and D. Kavanagh (eds), *The British General Election of 1992*, Basingstoke, Macmillan.
- — (1997), 'Appendix 2: The Results Analysed', in D. Butler and D. Kavanagh (eds), *The British General Election of 1997*, Basingstoke, Macmillan.
- Droop, H. R. (1871), 'On the Political and Social Effects of Different Method of Electing Representatives', ch. 3 in *Papers, 1863–70*, London, Juridical Society, 469–507.
- Duverger, M. (1954), *Political Parties: Their Organization and Activity in the Modern State*, New York, Wiley.
- Evans, G., Curtice, J., and Norris, P. (1998), 'New Labour, New Tactical Voting? The Causes and Consequences of Tactical Voting in the 1997 General Election', in D. Denver, J. Fisher, P. Cowley, and C. Pattie (eds), *British Elections and Parties Review, Volume 8: The 1997 General Election*, London and Portland, OR, Frank Cass, 65–79.
- Fisher, S. D. (2001), 'Definition and Measurement of Tactical Voting: The Role of Rational Choice', CREST Working Paper 95, available at [www.crest.ox.ac.uk](http://www.crest.ox.ac.uk)
- Hoffman, D. T. (1982), 'A Model for Strategic Voting', *SIAM Journal of Applied Mathematics*, **42**, 751–61.
- Lijphart, A. (1994), *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies, 1945–1990*, Oxford, Oxford University Press.
- McKelvey, R., and Ordeshook, P. (1972), 'A General Theory of the Calculus of Voting', in J. F. Herndon, and J. L. Bernd (eds), *Mathematical Applications in Political Science*, vol. 6, Charlottesville, VA, University Press of Virginia.
- Morris, S., and Shin, H. S. (2001), 'Global Games: Theory and Applications', available at <http://www.econ.yale.edu/smorris/>
- Morton, R. H. (2001), *How Elections Rule American Politics*, Norton, forthcoming.
- Myatt, D. P. (2002a), 'Idiosyncrasy, Information and the Impact of Strategic Voting', University of Oxford Department of Economics Discussion Paper Series No. 94.
- (2002b), 'A New Theory of Strategic Voting', University of Oxford Department of Economics Discussion Paper Series No. 93.
- Fisher, S. D. (2002), 'Everything is Uncertain and Uncertainty is Everything: Strategic Voting in Simple Plurality Elections', University of Oxford Department of Economics Discussion Paper Series No. 115.
- Myerson, R. B., and Weber, R. J. (1993), 'A Theory of Voting Equilibria', *American Political Science Review*, **87**(1), 102–14.
- Norris, P., Curtice, J., Sanders, D., Scammell, M., and Semetko, H. A. (1999), *On Message: Communicating the Campaign*, London, Sage.
- Palfrey, T. R. (1989), 'A Mathematical Proof of Duverger's Law', in P. C. Ordeshook (ed.), *Models of Strategic Choice in Politics*, Ann Arbor, MI, University of Michigan Press, 69–92.
- Riker, W. H. (1982), 'The Two-party System and Duverger's Law: An Essay on the History of Political Science', *American Political Science Review*, **76**(4), 753–66.
- Ordeshook, P. C. (1968), 'A Theory of the Calculus of Voting', *American Political Science Review*, **62**, 25–42.
- Simon, H. A. (1954), 'Bandwagon and Underdog Effects and the Possibility of Equilibrium Predictions', *Public Opinion Quarterly*, **18**, 245–53.