



DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

**EVERYTHING IS UNCERTAIN AND UNCERTAINTY IS
EVERYTHING: STRATEGIC VOTING IN SIMPLE PLURALITY
ELECTIONS**

David P. Myatt and Stephen D. Fisher

Number 115

July 2002

Manor Road Building, Oxford OX1 3UQ

EVERYTHING IS UNCERTAIN AND UNCERTAINTY IS EVERYTHING: STRATEGIC VOTING IN SIMPLE PLURALITY ELECTIONS

DAVID P. MYATT AND STEPHEN D. FISHER

DEPARTMENT OF ECONOMICS AND NUFFIELD COLLEGE, UNIVERSITY OF OXFORD

david.myatt@economics.oxford.ac.uk and steve.fisher@nuffield.oxford.ac.uk

Oxford University Department of Economics Discussion Paper no. 115. July 2002.

ABSTRACT. Intuition tells us that strategic voting is most likely in marginal constituencies where the preferred party is a long way behind the second placed party. Some formal theories suggest there should be complete desertion of all but two candidates (Palfrey 1989), or additionally that the second and third will have similar vote shares (Cox 1997). Unfortunately, these theories fail to account for uncertainty over the strength of candidates. We present a model that allows for such uncertainty. It generates interesting and original comparative statics. All three approaches are tested against English voting data from 1987, 1992 and 1997. Our model fits the data; the standard intuition and Cox hypothesis do not. Thus formal theory can improve on intuition. But, this depends on the realization that voters are uncertain, and it is only uncertainty that matters for strategic voting.

1. INTRODUCTION

“In these matters the only certainty is that there is nothing certain.” *Pliny the Elder*

The debate on strategic voting dates back at least to Pliny the Younger (Farquharson, 1969, pp. 57–60). However, it is our contention that the above quote from his uncle also shows great insight into strategic voting decision making. Voters are uncertain of the distribution of support for the candidates and, as we show, strategic incentives are entirely determined by this uncertainty.

On a general level, strategic voting in a first-past-the-post system is often thought to be motivated by the desire to avoid a “wasted vote” (Schattschneider 1942). If a voter’s preferred candidate seems likely to come third or lower they could vote for whichever of the leading two contenders they prefer. This will be tempting if the voter has a strong preference between the leading candidates, and more so if they are relatively indifferent between their preferred candidate and their favorite of the leading pair. The further behind the preferred candidate, the more a sincere vote will appear to be wasted. Likewise, the closer the contest between the leaders the more an elector will feel they have the opportunity to influence the result by casting a strategic vote. So runs the

We thank David Firth, Anthony Heath, Paul Klemperer, Iain McLean, David Sanders, Peyton Young and Christopher Wallace for their helpful comments, but remain responsible for any errors.

informal argument about the structure of strategic incentives which we refer to as the “standard intuition.”

Formal theorists have argued along different lines. They consider the behavior of short-term instrumentally rational voters, i.e. people who are certain to vote and care only about who wins in their district at the current election.¹ A strategic vote is correspondingly defined as a vote for a candidate thought more likely to win than the preferred candidate, in order to best influence who wins. Thus far there is no disagreement with the standard intuition; divergence comes with the approach to the voting decision. The only circumstance in which a voter can influence the outcome is when the votes of all the other electors produce a tie or near-tie for the lead. In these situations one additional vote can, respectively, break a tie or force a tie. Only in these situations is a voter *pivotal* because only in these situations can a single vote affect who wins. In any other situation, a single vote has no effect. So, non-pivotal outcomes are of no interest to an instrumental voter, and the absolute probability of being in a pivotal situation becomes irrelevant. An instrumental voter considers the relative chances of different pairs of parties being involved in a tie (or near-tie) for the lead. In effect they take it for granted that they are pivotal when deciding who to vote for. Instead of asking, “What are the chances of being pivotal?” as supposed by standard intuition, they ask, “Who will I be pivotal between supposing I am pivotal?”

This observation is a central to the work of Hoffman (1982), Palfrey (1989) and Cox (1994), and it may be used to draw some basic inferences about utility-maximizing voting behavior. For instance, a vote for a third placed candidate can be optimal and is no more a “wasted vote” than any other. Also, it may be that second placed candidates suffer from strategic desertion as well as third placed candidates. Since these theoretical findings are logical possibilities rather than positive predictions, they have limited use. For a fuller description of the pattern of strategic voting we need to make assumptions about what voters know about the preferences and behavior of others.

Palfrey (1989), Cox (1994) and Myerson & Weber (1993) describe models in which there is a probability distribution of preferences which is commonly known by voters, and individual voter preferences are drawn independently from that distribution. The Law of Large Numbers and the assumption of statistical independence together imply that the level of support for the different candidates in the electorate will be almost exactly the same as that in the commonly known distribution, if the electorate is sufficiently large (more than a few hundred). Statistical theory also

¹McKelvey & Ordeshook (1972) built upon the analysis of Riker & Ordeshook (1968) to consider three candidate competitions in which electors have the option of abstention, as well as sincere and strategic voting. However, it remains the case that with costly voting one expects electors to abstain.

shows us that the probability of a pivotal event involving the leading two candidates is infinitely larger than the probability of a pivotal event involving any other pair of parties. So an instrumental voter is certain of being pivotal between the leading pair, if they are to be pivotal at all. Since all the voters in the model realize this, they all vote for whichever of the leading pair they prefer. Only two candidates receive any votes and strategic voting is complete.

Palfrey (1989) describes this result as a “mathematical proof” of Duverger’s Law. Of course, Duverger (1954) only claimed that plurality voting yielded a tendency toward two-party systems. However, the two-candidate prediction that Palfrey (1989) describes could be described as a “strict Duvergerian” outcome. Palfrey’s analysis was supplemented by Cox (1994), who highlighted the existence of non-Duvergerian equilibrium outcomes which involve a tie between the second and third placed candidates. His interpretation of these equilibria is that voters are unsure of the exact support of the trailing candidates and hence the identity of the closest challenger to the lead (Cox 1997). The existence of two types of (equilibrium) outcome led Cox to propose his bimodality hypothesis: The ratio of support for the second to the first loser (the SF ratio) in a district should either be close to zero (Duvergerian equilibrium) or one (non-Duvergerian equilibrium).

The Cox-Palfrey approach is built upon a critical assumption: voters know the population characteristics of the distribution from which preferences are independently drawn. Whilst they are uncertain about individual behavior, voters are virtually certain of the election outcome under sincere voting if the electorate is sufficiently large. Since any tie for the lead almost always involves the leading two candidates, it is therefore not surprising that strict Duvergerian outcomes are predicted. The absence of any real uncertainty also calls into question Cox’s (1994) interpretation of his non-Duvergerian equilibria. In three candidate plurality elections, such equilibria require exactly the correct proportion of second placed candidate supporters to switch votes in the “wrong” direction (towards the trailing candidate) in order to generate a tie or near-tie for second place. Whilst Cox’s (1997) interpretation involves people not-knowing who to desert strategically, these equilibria actually require very precise knowledge of the true underlying support levels for each candidate. Furthermore, such equilibria are dramatically unstable (Fey 1997), and hence there the theoretical support for them is not overwhelming.

These observations lead to our central argument. The Cox-Palfrey statistical independence assumption needs to be relaxed. It is this assumption that drives the stark two-candidate result and leads to an effective absence of any real uncertainty. For large constituencies, individual idiosyncrasy is averaged out and voters may successfully predict the outcome of the election with

near certainty. This feature is recognized by the leading contributors to the voting literature. For instance, Cox (1997) comments:

“A fourth condition necessary to generate pure local bipartism is that the identity of trailing and front-running candidates is common knowledge . . . If who trails is not common knowledge, then an extra degree of freedom is opened up in the model.”

Since we know that voters in mass elections are never certain of the results and differ in their expectations of the order of candidates (see for example Evans & Heath (1993) and Abramson et al. (1992)), our aim is to open up this degree of freedom.

To this end, we present a formal decision-theoretic model of three-candidate competition with voter uncertainty at both the individual and constituency level. Even with relatively small electorates, any individual level uncertainty is averaged out. It is *only* uncertainty over the constituency characteristics that determines the behavior of an instrumental voter. Hence the claim in the title that uncertainty is everything. The predicted outcome is non-Duvergerian in the sense that all three candidates each receive votes. But, it could also be described as Duvergerian since strategic voting is predominantly away from the third placed candidate and towards the leading two candidates, thereby providing a tendency to bipartism.

The degree of strategic desertion depends on the relative support for the different candidates in the constituency and the precision of the information available to voters. The relationship between the level of strategic voting and the constituency situation is similar, but not the same as that in the standard intuition. In both theories, voters are more likely to abandon a party the further it falls behind the leading two parties, but after controlling for this effect our formal uncertainty model predicts that strategic incentives increase, not decrease, with the margin of victory.

A striking criticism of formal rational choice theory is that it merely formalizes intuitive ideas (Green & Shapiro 1994). Here formal theory generates predictions that differ from standard intuition. Apart from the desire to understand strategic voting, this is an ideal opportunity to test formal theories against competing intuition to see whether formal theory teaches us anything about political behavior. We do this using data on voters in England from the 1987, 1992 and 1997 British Election Studies (BES). This is a prime example of three party competition under the simple plurality electoral system. The data are inconsistent with the standard intuition and the Cox (1994) bimodality hypothesis, but consistent with the pattern of strategic incentive structure of our uncertainty model.

The paper is organized as follows. The formal model is presented in the following section, which may be omitted on first reading. The predictions of the uncertainty model are compared with the alternative standard intuition and Cox bimodality hypotheses in Section 3. These hypotheses are tested using data and methodology presented in Section 4 and the results are presented in Section 5. We conclude in Section 6.

2. FORMAL THEORY

2.1. A Single Seat Plurality Election.

A single seat plurality election is contested by m candidates, indexed by $j \in \{1, 2, \dots, m\}$. Since our model is applied to English constituencies, we will often focus on the $m = 3$ case. There are $n + 1$ individuals in the electorate, indexed by $i \in \{0, 1, \dots, n\}$. We will study an instrumentally rational voter $i = 0$, who will in turn consider the voting behavior of the remaining n voters. Among these other n individuals, we use x_j to denote the number of votes cast for candidate j , and the vector $x = [x_j]_{j=1}^m \in X$ to denote the election outcome, where the outcome space is $X = \{x \in \mathbb{Z}_+^m \mid \sum_{j=1}^m x_j = n\}$. Each individual casts a single vote, and there are no abstentions. The candidate receiving the highest number of votes (determined by the vote of $i = 0$ as well as x) wins the election. Of course, a tied outcome is possible, where a number of candidates share a common and leading vote total. Following British electoral convention, such ties are broken at random.²

To focus on the decision faced by voter $i = 0$, we exogenously specify the behavior of the remaining n voters. They are summarized by the parameter $p \in \Delta$, where $\Delta = \{p \in \mathbb{R}_+^m \mid \sum_{j=1}^m p_j = 1\}$ is the $m - 1$ dimensional unit simplex. Thus p represents the *constituency support* for the m candidates. A randomly selected individual $i > 0$ votes for candidate j with probability $p_j \in [0, 1]$: This is the *idiosyncratic uncertainty* in the model. Conditional on p , voting decisions for the n individuals $i > 0$ are independent. It follows that (given p) x is a draw from the multinomial distribution with parameters p and n .

It remains to specify the preferences and beliefs of voter $i = 0$. She is short term instrumentally rational, caring about the outcome of the election process only insofar as it influences the winning candidate. Her von Neumann-Morgenstern utility for a win by candidate j is u_j . Without loss of generality, the candidates are ordered so that $u_1 > u_2 > \dots > u_m$. Crucially, the parameter p is unknown her. It follows that, even in a large electorate ($n \rightarrow \infty$), she is uncertain of the

²In the United Kingdom the returning officer has a casting vote in tied situations. By convention, the officer determines the winner by the toss of a coin.

election outcome x . Uncertainty over p is represented by the density $f(p)$. This is assumed to be continuous and strictly positive on the interior of Δ , and represents the *constituency uncertainty* in the model. Given her beliefs $f(p)$ and preferences $\{u_j\}$, voter $i = 0$ maximizes her expected utility.

We note two features of this specification. First, p represents the likely voting *intentions* of the wider electorate and not necessarily their preferences. Voter $i = 0$ is interested in the electorate's preferences only insofar as they affect their votes and hence the outcome x . We do not, therefore, assume that the n individuals $i > 0$ vote sincerely — hence the model's applicability may be wider than initial expectations would suggest. Second, the model does not suffer from statistical independence. This parameter p is unknown to $i = 0$, and hence from her perspective votes are interdependent: The decision of one individual reveals information about p , and hence about the likely decisions of others.

2.2. Election Outcomes and Constituency Uncertainty.

An instrumental voter must evaluate the probabilities of different election outcomes in X . We begin, therefore, by considering the nature of such probabilities and the relative importance of constituency and idiosyncratic uncertainty. Conditional on p , the outcome x is multinomial with parameters p and n . Since voter $i = 0$ does not know p , she must take expectations with respect to the density $f(p)$. Explicitly:

$$\Pr[x] = \int_{\Delta} \Pr[x | p] f(p) dp = \int_{\Delta} \frac{n!}{\prod_{j=1}^m x_j!} \prod_{j=1}^m p_j^{x_j} f(p) dp = \frac{n!}{\prod_{j=1}^m x_j!} \int_{\Delta} \prod_{j=1}^m p_j^{x_j} f(p) dp \quad (1)$$

As n grows large, it is clear that such probabilities vanish to zero. For an instrumental voter, however, it is not the *absolute* probability of election outcomes that is of importance. Instead, it is the *relative* probability of different pivotal events, since it is only in the context of such events that her vote has any effect — we make this logic explicit in Section 2.3. We consider, therefore, the asymptotic behavior of outcome probabilities as $n \rightarrow \infty$. Examining the integrand in the final expression of Equation 1, notice that the function $\prod_{j=1}^m p_j^{x_j}$ is maximized at $p = x/n$. For large n , all mass in this function congregates around this maximum. It follows that the density $f(p)$ is only relevant local to $p = x/n$. Asymptotically, therefore, it is valid to replace $f(p)$ with $f(x/n)$. To make this precise, we require an appropriate definition of “asymptotic equivalence.”

Definition 1. For functions $g(x, n) > 0$ and $h(x, n) > 0$ write $g(x, n) \underset{n \rightarrow \infty}{\approx} h(x, n)$ whenever:

$$\lim_{n \rightarrow \infty} \max_{x \in X} \left\{ \frac{g(x, n)}{h(x, n)} \right\} = \lim_{n \rightarrow \infty} \min_{x \in X} \left\{ \frac{g(x, n)}{h(x, n)} \right\} = 1$$

If $g(x, n) \underset{n \rightarrow \infty}{\approx} h(x, n)$ then they are asymptotically equivalent in a large electorate.

In other words, we require $g/h \rightarrow 1$ as $n \rightarrow \infty$ irrespective of the choice of x along this sequence. This means that the convergence is uniform. Armed with this definition, we may now evaluate the asymptotic properties of $\Pr[x]$.

Proposition 1. *In a large electorate, $\Pr[x]$ is proportional to $f(x/n)$. Formally:*

$$\Pr[x] \underset{n \rightarrow \infty}{\approx} f\left(\frac{x}{n}\right) \frac{n!}{\prod_{j=1}^m x_j!} \int_{\Delta} \prod_{j=1}^m p_j^{x_j} dp = \frac{1}{n^{m-1}} f\left(\frac{x}{n}\right) \Rightarrow \frac{\Pr[x]}{\Pr[\tilde{x}]} \underset{n \rightarrow \infty}{\approx} \frac{f(x/n)}{f(\tilde{x}/n)}$$

so that, in a large electorate, likelihood ratios are determined entirely by the density $f(p)$.

Proof. This follows from the intuition given previously. The term $\prod_{j=1}^m p_j^{x_j}$ has a “spiked” maximum at x/n , and the density $f(p)$ is only of importance local to this value, so that $f(p)$ may be replaced by $f(x/n)$ in the limit. The equality follows from evaluation of the integral, which on inspection is the kernel of a Dirichlet distribution. The final claim follows as a corollary. The formal proof is relegated to Appendix A. □

Proposition 1 says that, in large electorates, the fraction of votes cast for each candidate will be close to x/n only when $p = x/n$, and hence the relative probability of observing such an outcome is determined by the density $f(x/n)$. This has two important consequences. First, it demonstrates that only constituency uncertainty (i.e. $f(p)$) matters: Idiosyncratic uncertainty is averaged out in a large electorate, and so constituency uncertainty is the critical ingredient. Second, the likelihood ratios of different outcomes are finite. This is not the case when p is known.

2.3. Voting Behavior.

We now consider the behavior of voter $i = 0$. She may only influence the identity of the winner in the context of a pivotal event: There must be a *tie* or *near tie* for the lead. When there is a tie,, a single additional vote may determine the winner. A near tie is a situation in which there is a winning margin of a single vote, and a single additional vote may create a tie and prevent an outright win by the leader. The pivotal events for a three candidate election (the focus of all subsequent analysis) are exhaustively listed in Table 1.

Since $u_1 > u_2 > u_3$ by assumption, voter $i = 0$ will never find it optimal to vote for candidate 3. Her choice, therefore, is a binary one, between candidates 1 and 2. In Table 1 we have listed the expected payoff to a vote for either candidate under each pivotal event, as well as the difference in

Event	$u(\text{Vote 1})$	$u(\text{Vote 2})$	$u(\text{Vote 1}) - u(\text{Vote 2})$	Description	
$x_1 = x_2 = x_3$	u_1	u_2	$(u_1 - u_3) - (u_2 - u_3)$	Three Way Tie	
$x_1 = x_2 = x_3 + 1$	u_1	u_2	$(u_1 - u_3) - (u_2 - u_3)$	Near Three Way Tie	
$x_1 = x_3 = x_2 + 1$	u_1	$\frac{u_1+u_2+u_3}{2}$	$\frac{u_1-u_2-u_3}{2}$	Near Three Way Tie	
$x_2 = x_3 = x_1 + 1$	$\frac{u_1+u_2+u_3}{2}$	u_2	$\frac{u_1+u_3-u_2}{2}$	Near Three Way Tie	
$x_1 = x_2 > x_3 + 1$	u_1	u_2	$(u_1 - u_3) - (u_2 - u_3)$	Two Way Tie	*
$x_1 = x_3 > x_2 + 1$	u_1	$\frac{u_1+u_3}{2}$	$\frac{u_1-u_3}{2}$	Two Way Tie	*
$x_2 = x_3 > x_1 + 1$	$\frac{u_2+u_3}{2}$	u_2	$-\frac{u_2-u_3}{2}$	Two Way Tie	*
$x_1 - 1 = x_2 = x_3$	u_1	$\frac{u_1+u_2}{2}$	$\frac{(u_1-u_3)-(u_2-u_3)}{2}$	Near Three Way Tie	
$x_1 - 1 = x_2 > x_3$	u_1	$\frac{u_1+u_2}{2}$	$\frac{(u_1-u_3)-(u_2-u_3)}{2}$	Near Two Way Tie	*
$x_1 - 1 = x_3 > x_2$	u_1	u_1	0	Near Two Way Tie	
$x_2 - 1 = x_1 = x_3$	$\frac{u_1+u_2}{2}$	u_2	$\frac{(u_1-u_3)-(u_2-u_3)}{2}$	Near Three Way Tie	
$x_2 - 1 = x_1 > x_3$	$\frac{u_1+u_2}{2}$	u_2	$\frac{(u_1-u_3)-(u_2-u_3)}{2}$	Near Two Way Tie	*
$x_2 - 1 = x_3 > x_1$	u_2	u_2	0	Near Two Way Tie	
$x_3 - 1 = x_1 = x_2$	$\frac{u_1+u_3}{2}$	$\frac{u_2+u_3}{2}$	$\frac{(u_1-u_3)-(u_2-u_3)}{2}$	Near Three Way Tie	
$x_3 - 1 = x_1 > x_2$	$\frac{u_1+u_3}{2}$	u_3	$\frac{u_1-u_3}{2}$	Near Two Way Tie	*
$x_3 - 1 = x_2 > x_1$	u_3	$\frac{u_2+u_3}{2}$	$-\frac{u_2-u_3}{2}$	Near Two Way Tie	*

TABLE 1. Pivotal Events and Voter Payoffs

payoffs between the two candidates. Voter $i = 0$ will calculate the probabilities of the various pivotal events, and will support candidate 1 whenever the expected difference in payoffs is positive, and support candidate 2 otherwise.

Although the number of potential pivotal events is large, only a subset are relevant. Examining Table 1, the near two-way tie $x_1 - 1 = x_3 > x_2$ involves a clear win for candidate 1 unless voter $i = 0$ were to vote for candidate 3. It has no relevance for the tradeoff between a vote for candidate 1 and a vote for candidate 2, and hence is an irrelevant pivotal event. Similar logic applies to the event $x_2 - 1 = x_3 > x_1$. To simplify consideration of the remaining pivotal events, we allow the electorate to grow large ($n \rightarrow \infty$). Naturally, the probability of a pivotal event will vanish to zero. As we have argued, however, it is only the *relative* likelihood of different pivotal events that is important. The asymptotic behavior of pivotal probabilities is recorded in the following lemma.

Lemma 1. *Two-way and near two-way ties are asymptotically equivalent; for instance:*

$$\Pr[x_1 - 1 = x_2 > x_3] \underset{n \rightarrow \infty}{\approx} \frac{\int_{1/3}^{1/2} f(z, z, 1 - 2z) dz}{6n} \underset{n \rightarrow \infty}{\approx} \Pr[x_1 = x_2 > x_3 + 1] \quad (2)$$

with similar expressions available for $\Pr[x_1 = x_3 > x_2]$ and $\Pr[x_2 = x_3 > x_1]$. Three way and near three-way ties are asymptotically equivalent, and satisfy:

$$\Pr[x_1 = x_2 = x_3 + 1] \underset{n \rightarrow \infty}{\approx} \frac{1}{n^2} f\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \underset{n \rightarrow \infty}{\approx} \Pr[x_1 = x_2 = x_3] \quad (3)$$

where the first equivalence applies for $(n - 1)/3 \in \mathbb{N}$ and the second for $n/3 \in \mathbb{N}$.

Proof. Equation 2 follows from the application of Proposition 1, and summation over outcomes x that constitute a two way tie. This summation converges to the stated integral, which ranges over constituency support levels p satisfying $p_1 = p_2 > p_3$. Equation 3 is a direct application of Proposition 1. The qualification referring to n ensures that the events under consideration actually exist. Appendix A contains a formal proof. \square

Lemma 1 confirms that the probability of a pivotal outcome vanishes to zero as $n \rightarrow \infty$. Notice, however, that the probability of a three-way tie vanishes at rate n^{-2} , whereas the probability of a two-way tie vanishes at rate n^{-1} , and so three-way and near three-way ties are asymptotically irrelevant to voter $i = 0$. The remaining relevant pivotal events are highlighted (“*”) in Table 1. The payoff differences between a vote for 1 and a vote for 2 depend only on the preference for candidates 1 and 2 relative to candidate 3. It follows that that the optimal voting rule may be couched in terms of these relative payoffs.

Proposition 2. *Assuming without loss of generality that an indifferent voter $i = 0$ casts her vote in favor of candidate 1, the optimal voting rule is asymptotically equivalent to:*

$$\text{Vote 1} \Leftrightarrow \underbrace{\log \frac{u_1 - u_3}{u_2 - u_3}}_{\text{Relative Preference}} \geq \log \frac{2 \Pr[x_1 = x_2 > x_3] + \Pr[x_2 = x_3 > x_1]}{2 \Pr[x_1 = x_2 > x_3] + \Pr[x_1 = x_3 > x_2]} \quad (4)$$

$$\underset{n \rightarrow \infty}{\approx} \log \underbrace{\frac{p_{23} + 2p_{12}}{p_{13} + 2p_{12}}}_{\text{Strategic Incentive}} \equiv \lambda \quad (5)$$

where $p_{12} = \int_{1/3}^{1/2} f(z, z, 1 - 2z) dz$, $p_{13} = \int_{1/3}^{1/2} f(z, 1 - 2z, z) dz$ and $p_{23} = \int_{1/3}^{1/2} f(1 - 2z, z, z) dz$.

Proof. To show this, assemble the asterisked components of Table 1 to obtain the optimal voting criterion. This applies asymptotically, since the three way ties have been removed. Impose the asymptotic equivalence of two-ties and near two-way ties from Lemma 1, and re-arrange to obtain Equation 4. Use Equation 2 of Lemma 1 to obtain Equation 5. \square

Voter $i = 0$ balances her relative preference for candidates 1 and 2 against the relative likelihood of influencing the election outcome. For instance, a vote for candidate 1 rather than candidate 2 yields a payoff gain of $u_1 - u_3$ whenever there is a tie between candidates 1 and 3. Furthermore, whenever candidates 1 and 2 are tied, there is a gain of $u_1 - u_2 = (u_1 - u_3) - (u_2 - u_3)$. This latter tie carries the coefficient 2, since the vote switch has twice the effect: The loss of a vote for candidate

2 and the gain of a vote for candidate 1. Assembling such effects yields the λ as an appropriate measure of the strategic incentive. This can be difficult to interpret, since it takes values on the entire (positive and negative) real line. An alternative, and intuitive, representation is as follows.

Definition 2. When $\lambda \geq 0$, define the intensity of the strategic incentive as $\Lambda = 1 - e^{-\lambda}$.

This has a number of attractive properties. First, it is increasing in λ , satisfying $\Lambda = 0$ when $\lambda = 0$ and $\Lambda \rightarrow 1$ as $\lambda \rightarrow \infty$. Thus it can be interpreted as the proportion of the maximum possible strategic incentive. Second, it measures how strong a voter's allegiance to her preferred candidate must be in order to avoid a strategic vote. To see this, use Equation 5 to confirm that $i = 0$ will vote for her preferred candidate when:

$$\frac{u_1 - u_3}{u_2 - u_3} \geq e^\lambda \quad \Leftrightarrow \quad e^{-\lambda} \geq \frac{u_2 - u_3}{u_1 - u_3} \quad \Leftrightarrow \quad \frac{u_1 - u_2}{u_1 - u_3} \geq 1 - e^{-\lambda} = \Lambda$$

When Λ is large, $i = 0$ will vote strategically unless her preference for her preferred candidate relative to her second choice ($u_1 - u_2$) is large in comparison to her preference for her preferred candidate relative to the disliked candidate ($u_1 - u_3$). Thus strategic voting is more likely when when a voter is relatively indifferent between her first and second choices (so that u_1 is close to u_2) and when she has a strong dislike of the third candidate (so that $u_1 - u_3$ is very large). These predictions are familiar, and coincide with those posited in the empirical literature (Heath et al. 1991).

Importantly, the strategic incentive λ is *finite*. With uncertainty of the constituency-wide support for the various candidates, the incentive to vote strategically remains limited, even in an unboundedly large electorate. This opens the possibility for multi-candidate support. Moreover, the incentive varies with the structure of a voter's beliefs, enabling further investigation of the determinants of strategic voting.

2.4. Modeling Constituency Uncertainty.

Proposition 2 justifies a subsequent focus on a voter's beliefs $f(p)$, and an approach similar to that adopted by McKelvey & Ordeshook (1972) and Hoffman (1982). These authors define beliefs directly over X , using a continuous approximating density. The analogue here is the density $f(p)$. McKelvey & Ordeshook (1972) specify a uniform distribution over X , whereas Hoffman (1982) considers a density with a Gaussian kernel. Unfortunately, neither of these approaches yields clear and testable comparative statics.

We suggest the use of the Dirichlet distribution as an appropriate specification for a voter's beliefs. It is well known, tractable, and (as we will show) offers interesting and specific comparative static predictions. Most importantly, however, it is conjugate to the multinomial distribution, and hence arises naturally from a procedure in which a voter samples the voting decisions of others in the electorate. To see this, suppose that a voter is initially ignorant, beginning with a uniform prior over Δ . She then observes the voting intentions of random sample of s other voters. Write y for a vector with elements y_j for $j \in \{1, 2, \dots, m\}$ denoting the number of sampled individuals who support candidate j , so that $\sum_{j=1}^m y_j = s$. Bayesian updating following the observation y yield posterior beliefs:

$$f(p|y) \propto \Pr[y|p]f(p) = \frac{s!}{\prod_{j=1}^m y_j!} \prod_{j=1}^m p_j^{y_j} \propto \prod_{j=1}^m p_j^{y_j} \quad (6)$$

This last expression is the kernel of a Dirichlet distribution with parameters $\{y_j + 1\}_{j=1}^m$. For convenience, define $\pi_j = y_j/s$, the fraction of the sample of voters who support candidate j , and assemble into the vector $\pi \in \Delta$. With this formulation we may write:

$$f(p|y) \propto \tilde{f}(p) = \left[\prod_{j=1}^m p_j^{\pi_j} \right]^s$$

Examining $\tilde{f}(p)$, notice that $\pi = \arg \max \tilde{f}(p)$. It follows that $\pi \in \Delta$ is the *modal belief* for the focal voter, and s indexes the *precision* of beliefs around this mode. In summary, an initially ignorant voter (interpreted as a uniform prior) who bases her opinions on a random sample of voting intentions will hold Dirichlet beliefs where π is the most likely constituency support level in her eyes, and s is the quantity of information available to her. We may use the Dirichlet in conjunction with Proposition 2. This yields an explicit, closed form strategic incentive variable in the context of a three candidate election.

Proposition 3. *For Dirichlet beliefs with $m = 3$, the strategic incentive satisfies:*

$$\lambda = \log \frac{\int_{1/3}^{1/2} [2g(z, \pi_3)^s + g(z, \pi_1)^s] dz}{\int_{1/3}^{1/2} [2g(z, \pi_3)^s + g(z, \pi_2)^s] dz} \quad \text{where} \quad g(z, \pi) = z^{1-\pi}(1-2z)^\pi \quad (7)$$

Evaluating the integrals, this may be expressed as:

$$\lambda = \log \frac{B_{1/3}(1 + \pi_3 s, 1 + (1 - \pi_3)s) + 2^{(\pi_1 - \pi_3)s-1} B_{1/3}(1 + \pi_1 s, 1 + (1 - \pi_1)s)}{B_{1/3}(1 + \pi_3 s, 1 + (1 - \pi_3)s) + 2^{(\pi_2 - \pi_3)s-1} B_{1/3}(1 + \pi_2 s, 1 + (1 - \pi_2)s)} \quad (8)$$

where $B_{1/3}(a, b)$ is the incomplete Beta function evaluated at $1/3$ with parameters a and b . $\lambda > 0 \Leftrightarrow \pi_2 > \pi_3$. Fixing π_3 , λ is increasing in $\pi_2 - \pi_1$. Fixing π_2 , λ is increasing in $\pi_3 - \pi_1$.

Proof. See Appendix B. □

The Dirichlet belief system yields immediate predictions. First, it is unnecessary for a voter to expect her preferred candidate to trail in third place before the strategic incentive is positive: $\pi_2 > \pi_1 > \pi_3$ yields a positive incentive to vote strategically. This means that strategic desertion can damage parties placed second in the constituency. Furthermore, lowering the support of the preferred candidate always increases the incentive to vote strategically. Proposition 3 does not, however, say anything about comparative statics as support moves from candidate 2 to candidate 3. We explore this issue in Section 3.1.

3. COMPARISON WITH ALTERNATIVE HYPOTHESES

3.1. Marginality, Distance from Contention and the Standard Intuition.

The empirical literature displays some consensus on the determinants of strategic voting. It is thought to be more attractive to voters who are relatively indifferent between their first and second preference parties, but strongly prefer their second over their third choice party (Heath et al. 1991). Typically it is those who support a candidate likely to come third or lower who feel the need to make a strategic switch to, “avoid wasting their vote”. Since people are never sure that their preferred candidate is in third place, the more a voter feels their candidate is out of the running, the more willing they should be to switch strategically. Thus strategic voting is expected to increase as the difference in support between the second placed candidate and the first preference candidate increases. This statistic is known as the “distance from contention” and is well known as a strong predictor of strategic voting.³ Finally, strategic voting should intuitively be more likely when the chances of influencing the result are greatest. Thus strategic voting should be greater in more marginal constituencies because the absolute probability of a tie is greater when the “margin of victory”—the gap between the leading candidates—is small. This is a commonly held proposition (see, for instance, Cain (1978), Niemi, Whitten & Franklin (1992), Evans (1994), Fieldhouse, Pattie & Johnston (1996) and Cox (1997)). It has even been used to infer strategic voting from aggregate election results (Curtice & Steed (1988, 1992, 1997)). However, the evidence for it has always been very weak. No one has found evidence for a marginality effect that is not specific to supporters of a particular party. Nevertheless, the standard intuition is a popular hypothesis which comprises four interrelated propositions.

³The variable is such as strong predictor that it serves as a basis of construct validity tests for different measures of tactical voting (Franklin *et al* 1992, 1993, 1994 and Evans & Heath 1993, 1994).

Hypothesis 1 (The Standard Intuition). *For supporters of third placed candidates, strategic voting should decrease with the relative strength of preference for the favorite over the second choice candidate, and increase with the relative strength of preference for the second over the third choice candidate. Strategic voting should increase with the distance from contention, and decrease with the margin of victory.*

There is a great deal of similarity between the predictions of the standard intuition and the corresponding comparative statics from the uncertainty model presented in the previous section. Rewriting the first term of Proposition 2,

$$\log \frac{u_1 - u_3}{u_2 - u_3} = -\log \left\{ \frac{u_1 - u_2}{u_2 - u_3} + 1 \right\}$$

we can see that strategic voting increases with the relative strength of preference for the favorite over the second choice ($u_1 - u_2$) and decreases with the relative strength of preference for the second over the third choice candidate ($u_2 - u_3$).⁴

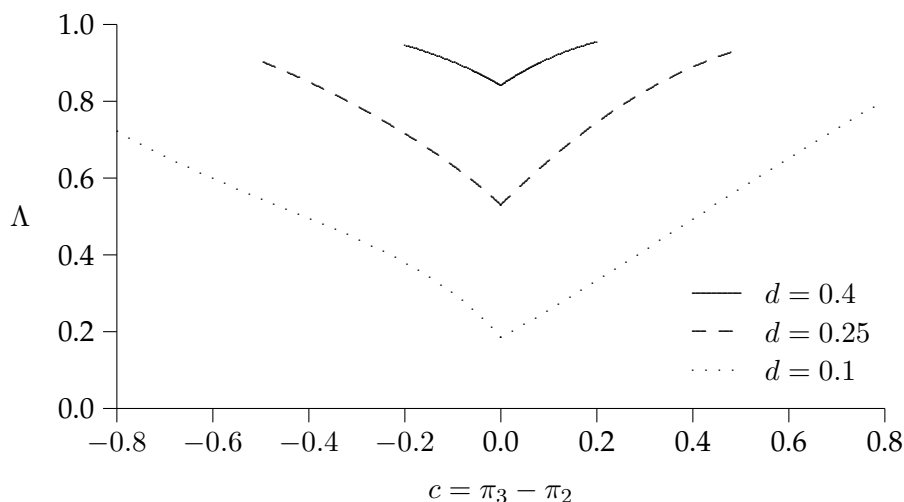
A similar claim can be made regarding distance from contention, but not marginality. Consider the situation where the preferred candidate trails in third place, so either $\pi_2 > \pi_3 > \pi_1$ or $\pi_3 > \pi_2 > \pi_1$.⁵ The distance from contention is then $d = \min\{\pi_2, \pi_3\} - \pi_1$ and the margin of victory is $|\pi_3 - \pi_2|$. Since the margin of victory does not distinguish which of the second and third preference candidates is ahead, let c be the winning margin for the most disliked candidate. So $c = \pi_3 - \pi_2$, and $|c|$ is then the margin of victory. The effect of the distance from contention d and the winning margin for the most disliked candidate c on the intensity of the strategic incentive $\Lambda = 1 - e^{-\lambda}$ (see Definition 2) is illustrated in Figure 1. For constant c the incentive to vote strategically increases with d , the distance from contention. This accords with the standard intuition. However, for any given value of d , the strategic incentive increases as c moves away from zero. So, holding the distance from contention constant, the strategic incentive increases with the margin of victory. This is the exact opposite of the standard intuition prediction. More formally we can state the following.

Proposition 4. *For $\pi_1 < \min\{\pi_2, \pi_3\}$, the strategic incentive λ is increasing in both the distance from contention d and the margin of victory $|c|$. Precisely, $\partial\lambda/\partial c > 0 \Leftrightarrow c > 0$.*

Proof. See Appendix B. □

⁴Note from subsection 2.1 that u_j is the utility for a win by candidate j and candidates are ordered so that $u_1 > u_2 > u_3$.

⁵Note from subsection 2.4 that π_j is the proportion of the vote for candidate j in a hypothetical poll.



Fixing the distance from contention d , this chart shows the effect of the margin of victory for the most disliked candidate c on the intensity of the strategic incentive $\Lambda = 1 - e^{-\lambda}$ when the preferred candidate is in third place. The precision level is fixed at $s = 10$.

FIGURE 1. Effect of Marginality on Strategic Incentives in the Uncertainty Model.

What explains the difference between the standard intuition and our formal model? The premise of the informal argument is that voters are more willing to vote strategically when the chances of influencing the result are greater. Since the probability of being pivotal increases with marginality, strategic voting should decrease with the margin of victory. In the formal analysis it is not the *absolute* probability of a tie that is relevant to an instrumental voter, but the *relative* probability of different pairs of parties being involved in a tie for the lead. To make it worthwhile voting sincerely, a third placed candidate supporter must think there is a sufficient chance that any tie for the lead will involve their preferred candidate. Holding the distance from contention constant, an increase in the margin of victory reduces the *absolute* probability of a tie for the lead, but the probability that a tie involves the leading two candidates increases *relative* to the probability that a tie involves the preferred candidate. So the incentive to vote strategically increases with the margin of victory after controlling for the distance from contention. This argument allows us to state a hypothesis that differs from the standard intuition.

Hypothesis 2 (Voter Uncertainty). *For supporters of third placed candidates, strategic voting should decrease with the relative strength of preference for the favorite over the second choice candidate, and increase with the relative strength of preference for the second over the third choice candidate. Strategic voting should increase with both the distance from contention and margin of victory, or alternatively and more precisely, with the strategic incentive λ defined in Proposition 3.*

3.2. Precision of Information and the Cox-Palfrey Framework.

Our central innovation is the introduction of constituency uncertainty. This is parameterized by s , which is effectively the sample size for a hypothetical poll of voting intentions, and we can examine the strategic incentive as beliefs become more or less precise. In particular, it is interesting to note that for third placed candidate supporters, the strategic incentive is extremely large when beliefs are very precise. The opposite holds when the preferred candidate is expected to run second. This is unsurprising, for when s is large the voter is sure of the constituency situation and the identity of the leading two candidates. More formally we can state.

Proposition 5. *If $\pi_1 < \min\{\pi_2, \pi_3\}$ then $\lim_{s \uparrow \infty} \lambda = \infty$. If $\pi_3 < \pi_1 < \pi_2$ then $\lim_{s \uparrow \infty} \lambda = 0$.*

Proof. See Appendix B. □

This Proposition allows an interpretation of Palfrey's (1989) "mathematical proof" of Duverger's Law. He assumed that the decisions of voters were drawn independently from a commonly known distribution. In the context of the present theory, this is equivalent to allowing s to tend to infinity. If a voter's preferred candidate is expected to trail in third place then the strategic incentive is infinite and always vote strategically. Similarly, if a preferred candidate is expected to finish first or second, then there will be no strategic incentive. The conclusion is that all supporters of the third placed party will vote strategically, leading to a strictly Duvergerian two-candidate outcome. Palfrey's prediction, therefore, is that only two candidates will receive votes when the electorate is perfectly informed about candidate support levels.

Close inspection of Proposition 5, however, reveals that this analysis is contingent on there being a strict gap between the second and third placed candidates. If there is a tie in support between the second and third placed candidates (for instance, when $\pi_1 = \pi_2 < \pi_3$) then Proposition 5 no longer applies. This observation is the foundation for Cox's (1994) "non-Duvergerian" equilibrium. Using a statistically independent framework, he demonstrated the existence of an equilibrium in which the two candidates are tied for second place. Despite the mathematical possibility of such an outcome, it is extremely suspect on theoretical grounds for three reasons. First, it requires voters to have exact knowledge of the support for the different candidates. Second, unless the true support for the trailing candidates is perfectly balanced (i.e. $p_1 = p_2 < p_3$), it requires a *precise* fraction of supporters of a more popular candidate to switch in the "wrong" direction toward a less popular candidate. Yet, the chances of a tie for second place by accident are extremely slim, and there is no reason why voters with perfect information should switch in the wrong direction. Third, the equilibrium generated is highly unstable. Any small deviation would result in a very large strategic incentive for supporters of the trailing candidate (Proposition 5), leading back to

a Duvergerian outcome. Thus the Cox (1994) non-Duvergerian equilibrium is highly unstable (Fey 1997) and should not be expected in real life.⁶

The intuition offered by Cox (1997, p. 86) contradicts the formal mechanisms involved in his non-Duvergerian equilibria. He considered the result of the Ross and Cromarty constituency in the 1970 British General Election. A split in the vote between the Liberal and Labour candidates permitted a Conservative win. His non-Duvergerian interpretation was that “it was not clear who was in third and who in second.” But the non-Duvergerian equilibria in Cox’s model require exact knowledge of the true support for candidates. Irrespective of any theoretical dissatisfaction, however, Cox’s classification of equilibria does generate a hypothesis that may be taken to the data. His claim is that constituencies should be clustered in two groups: those in which the top two parties receive nearly all the votes between them and those in which the second and third placed parties have a very similar share of the vote. Cox devised a test of this hypothesis (Cox 1997, pp. 86–9) using the SF ratio, the ratio of the share of the voter for the Second loser to that of the First loser (third and second placed candidates respectively in a simple plurality system). If Cox is right, the distribution of the SF ratio across electoral districts should have two modes corresponding to the two types of equilibrium. This is the bimodality hypothesis and can be stated as follows.

Hypothesis 3 (Cox Bimodality). *The SF ratio should tend to either 0, because the third and lower placed candidates are strategically deserted, or to 1 because it is unclear which of the second and third placed candidates should be deserted.*

4. DATA AND METHODOLOGY

4.1. Data.

The British Election Studies (BES) for 1987, 1992 and 1997 provide high quality post-election survey data with sufficient questions to study strategic voting in depth (Heath, Jowell & Curtice 1995, Heath et al. 1996, Heath et al. 2000). Analysis is based on voters in England only because Scotland and Wales both have very strong nationalist parties and therefore genuine four party competition. Although there are minor parties in England, the same three major parties (Conservative, Labour and Liberal Democrat) stood at all three elections and took first, second and third places in all but a couple of constituencies (which have been excluded from the analysis). Minor

⁶Furthermore, in contrast to the claims of Cox (1994), non-Duvergerian equilibria are non-generic and depend critically on common knowledge of the constituency situation. If beliefs over the constituency situation are almost perfect, but differ just slightly from voter to voter, then non-Duvergerian equilibria of the kind Cox describes do not exist.

parties within England are ignored, partly because the voter uncertainty model does not provide for more than three parties, but mainly because one cannot tell from the BES data where minor parties lie in the party preference orderings of the respondents. Non-voters were also excluded because the hypotheses we are testing assume that people do vote.

4.2. Measurement of Strategic Voting.

Strategic voting is measured using responses to the following BES question.

- A. Which one of the reasons on this card comes closest to the main reason you voted for the party you chose?
1. I always vote that way
 2. I thought it was the best party
 3. I really preferred another party but it had no chance of winning in this constituency
 4. Other (write in)
 5. None of these/Don't know

Strategic voters identified by response option 3 to question A were asked a follow-up question.

- B. Which was the party you really preferred?

Respondents who gave tactical reasons for their vote in option 4 to question A were also coded as strategic if this was consistent with both their declared voting behavior and other questions relating to their order of preference for the parties. By definition strategic voters do not vote for their preferred party, so if there was any indication that they did so then the respondent was not coded as strategic. This was done using both question B and the "strength-of-feeling scores" (or "approval ratings") for the parties. The strength-of-feeling score for a party is the response coding (1 to 5) from the following question about the party.

- C. Please choose a phrase from this card to say how you feel about the (Conservative Party / Labour Party / Liberal Democrats / ...)
1. Strongly in favour
 2. In favour
 3. Neither in favour nor against
 4. Against
 5. Strongly against

When the respondent is a strategic voter the preferred party is provided by the response to question B, or is imputed from the strength-of-feeling scores. For non-strategic voters the party voted for is the first preference party, unless there is a clear indication otherwise on the strength-of-feeling scores. The second preference party is defined, for all respondents, as the party with the best strength-of-feeling score that is not the first preference party. The third preference party is similarly defined. Sometimes there is a tie for second preference on the strength-of-feeling scores.

This is either decided according to who the respondent said they would vote for if they had a second vote, or, in a small minority of cases (6.2%), the identity of the second choice is chosen so as to prejudice the test against the voter uncertainty model.⁷ Results under other coding schemes are similar.

The level of strategic voting in England according to this measure was 5.0% in 1987, 7.7% in 1992 and 8.5% in 1997. But it is more illuminating to measure strategic voting as a proportion of those voters who actually faced a strategic decision — the “risk population.” Blais & Nadeau (1996) also suggested the idea of a risk population to aid identification of strategic voting, but it is most valuable for analysis. Since variables such as distance from contention and the margin of victory are strongly correlated with the size of the third party vote, any association between them and strategic voting within the population of all voters reflects differences in the size of the risk population rather than salient aspects of voter decision making among those considering a strategic vote. This problem can be overcome by limiting analysis to the risk population. Ideally members of the risk population would be identified with reference to their prior expectations of the result. Unfortunately, this information is not available from the BES. So the risk population is pragmatically defined as all those voters whose preferred party came third or lower in the constituency at the election under investigation, at the previous election, or in a poll estimate of the election result.

The risk population so defined is too large for two reasons. Firstly, since any of the three result sets can be used to establish third placed party supporter status there is a greater chance of being in the risk population than if only one or all sets of results were required. This feature of the definition is especially relevant in 1997 when there was substantial change in the placing of parties within constituencies between 1992 and 1997. Nonetheless, it is important to retain consistency in the method of analysis across elections. Secondly, as Blais & Nadeau (1996, Endnote 7, p.50) claim, using electoral results rather than direct measures of expectations systematically inflates the population considered at risk of strategic voting because individuals overestimate the winning chances of their preferred party. The 1997 British Election Campaign Study (Norris et al. 1999) shows that 38% of those whose preferred party came third had expected it to come first or second. Given that only 13% of those whose preferred party came first actually expected another to win, there is clear evidence for inflated expectations for the favorite. However, more than four times as

⁷In such cases, for a strategic voter the second choice is the party with the lowest vote share of the two. For non-strategic voters, the second preference party is the party with the highest vote share of the two parties. The use of this coding provides a “worst case scenario” test of the voter uncertainty model, but is of no consequence for the test of the Cox model or standard intuition hypotheses.

many seats changed hands in 1997 than at the 1992 or 1987 elections, and in several constituencies parties moved from third to first place or *vice versa*. So mistaken expectations, including those related to overestimation of the favorite, are likely to be greater in 1997 than previously. Also roughly half of the risk population indicated that they expected their preferred party to come third or lower which, given item non response of 17.5%, shows that expectations of the chances of the preferred party are still fairly accurate.

Despite the problem of an imperfect pragmatic definition, the risk population still includes most people who could reasonably have voted strategically and relatively few for whom it would be unreasonable to do so. Among major party supporters in the risk population strategic voting was at 12.6% in 1987, 21.4% in 1992 and 24.4% in 1997. So strategic voting is a substantial phenomenon for the population for whom it is a relevant option. Finally, it is worth noting that although the Liberals came third nationally in each election, they came first or second in roughly half of the constituencies. The risk population is composed of 4.9% Conservative, 53.5% Liberal and 41.6% Labour supporters on average. All parties are represented in the risk population, not just the third party nationally.

4.3. Explanatory Variables.

Measurement of the relative strength of preference is based on the strength-of-feeling scores from question C. For ease of reference, the relative strength of preference for the favorite over the second choice is the *first preference gap*, while the *second preference gap* is the relative strength of preference for the second over the third choice party. The preference gap between two parties is the difference in the strength-of-feeling scores for the two parties. The *distance from contention* is the difference in the constituency share of the vote between the preferred party and the party placed second in the constituency. The *margin of victory* is the difference between the share of the vote for the winning and second placed parties in the constituency. For the voter uncertainty model, calculation of the *strategic incentive* variable from Equation 7 is described in Appendix C. The distance from contention, margin of victory and strategic incentive variable are all measured on the percentage share of the vote in the election under consideration, however analyses in which these variables are measured using previous election results or poll estimates of the constituency results produce similar results.

5. HYPOTHESIS TESTING

5.1. The Standard Intuition.

Variable	1987			1992			1997		
	B	s.e.	p	B	s.e.	p	B	s.e.	p
First preference gap (base=0)			0.00			0.00			0.00
First gap = 1	-0.90	0.31	0.00	-1.18	0.26	0.00	-1.07	0.25	0.00
First gap = 2	-1.81	0.50	0.00	-1.85	0.43	0.00	-2.42	0.67	0.00
First gap = 3 or more	-1.89	1.12	0.09	-7.42	12.28	0.55	-1.92	1.54	0.21
Second preference gap (base=0)			0.05			0.14			0.00
Second gap = 1	0.39	0.52	0.46	0.38	0.42	0.36	1.39	0.48	0.00
Second gap = 2	1.06	0.51	0.04	0.61	0.41	0.14	1.98	0.47	0.00
Third gap=3 or more	1.12	0.55	0.04	1.03	0.47	0.03	2.30	0.48	0.00
Distance from contention	0.07	0.02	0.00	0.09	0.02	0.00	0.05	0.01	0.00
Margin of victory	0.05	0.01	0.00	0.02	0.01	0.15	0.00	0.01	0.76
Constant	-3.75	0.69	0.00	-2.68	0.63	0.00	-2.58	0.58	0.00

(a) Hypothesis 1: The Standard Intuition

Variable	1987			1992			1997		
	B	s.e.	p	B	s.e.	p	B	s.e.	p
First preference gap (base=0)			0.00			0.00			0.00
First gap = 1	-0.89	0.31	0.00	-1.14	0.26	0.00	-1.05	0.25	0.00
First gap = 2	-1.77	0.49	0.00	-1.87	0.43	0.00	-2.44	0.67	0.00
First gap = 3 or more	-1.80	1.12	0.11	-7.37	12.32	0.55	-1.89	1.54	0.22
Second preference gap (base=0)			0.05			0.18			0.00
Second gap = 1	0.36	0.52	0.49	0.34	0.42	0.41	1.49	0.48	0.00
Second gap = 2	1.01	0.51	0.05	0.55	0.41	0.18	2.11	0.47	0.00
Third gap=3 or more	1.16	0.54	0.03	0.96	0.47	0.04	2.43	0.48	0.00
Strategic incentive (s=20)	0.45	0.14	0.00	0.68	0.11	0.00	0.56	0.13	0.00
Constant	-2.39	0.50	0.00	-1.82	0.42	0.00	-2.71	0.49	0.00

(b) Hypothesis 2: The Voter Uncertainty Model

Notes: B is the coefficient, s.e. is the standard error and p is the p-value. P-values where no coefficients are show are those from the Wald test of the variable. The sample from the BES includes major party supporters in England in the risk population (see text for further information). In 1987 N=603, in 1992 N=435 and in 1997 N=465.

TABLE 2. Logistic Regressions Testing Hypotheses 1 and 2.

The four comparative static predictions of Hypothesis 1 need to be analyzed simultaneously because of structural relationships between the independent variables. The relative strength of preference for the first over the second choice and that for the second over the third choice both depend on the strength-of-feeling score for the second choice party. Also, the distance from contention and margin of victory are variance dependent and *a priori* negatively correlated. Thus an effect of distance from contention on strategic voting can potentially produce the appearance of a marginality effect and *vice versa*.

Table 2(a) presents coefficients from logistic regression analyses of the decision to vote strategically that include all four explanatory variables suggested by the standard intuition. The coefficients

for the different variables are remarkably similar across the different elections although the p -values do vary. The relationship between strategic voting and the preference gaps is broadly as expected in each election.⁸ Also as predicted, the chances of voting strategically increase as the distance from contention increases. The characteristic proposition of the standard intuition that differs from the voter uncertainty model is that strategic voting should decrease as the margin of victory increases. However, the relationship observed in the data was the exact opposite in 1987, and in 1992 and 1997 there was no relationship to speak of. In a pooled analysis of the three elections we find similar results — strategic voting actually increases with the margin of victory and the relationship is even significant at the 10% level. The remaining three variables in the pooled analysis are statistically highly significant and in the direction expected, but the failure of the marginality hypothesis in the standard intuition is a crucial one.

5.2. The Voter Uncertainty Model.

A similar test to that for the standard intuition can be devised for Hypothesis 2. Here the predictor variables are the strength of preference for the first over the second choice party (first gap), the strength of preference for the second over the third choice (second gap) and the strategic incentive variable from Equation 7. The results are shown in Table 2(b). The first two explanatory variables are familiar from the test of the standard intuition. As before, strategic voting clearly decreases with the first preference gap and increases with the second preference gap. The association with the second gap is not significant for 1987 and 1992 and it is certainly weaker than the association with the first gap. However, the coefficients for each election show a monotonic decrease as the second gap increases and the result is significant in a pooled analysis of all three elections.

The strategic incentive from the voter uncertainty model was computed from Equation 7 for various different values of the precision of information variable. A range of values between 2 and 100 have been tested and the incentive has a statistically significant positive coefficient in each case. Table 2(b) gives the results when the precision of information is $s = 20$. This value was chosen because it is a reasonable precision level and the coefficient of the incentive variable seems to have a maximum near 20. The test is also robust to the coding of those with a tie for second and third choice. If the test is repeated with the opposite coding from that described in Section 4, the results are similar. Furthermore, the model is in fact more general than the standard intuition because it specifies a strategic incentive for all voters not just the members of the risk population, assuming

⁸The magnitude of the estimates and standard errors of the effects when a gap is “3 or more” can be explained by the fact that these are based on low numbers of respondents in these categories. If we make allowances for this term the patterns in the analysis are exactly as expected.

that we can make sense of a negative incentive. Testing the model on all major party supporters again produces similar results, although sensitivity to all variables, especially the incentive variable, is substantially reduced. This is unsurprising given the dominance in the voting population of those who support the first or second placed party in their constituency. These people have a negative strategic incentive which varies substantially in its strength, but they experience nearly no variance in the outcome since they nearly never vote strategically.

Perhaps most remarkable is the finding that a pooled analysis of all three elections including the strength of preference variables, the strategic incentive, and both the distance from contention and margin of victory, shows that the strategic incentive is a significant predictor while the distance from contention and margin of victory are insignificant. This is impressive considering the high structural correlation between the strategic incentive and both the distance for contention and the margin of victory. In contrast with both the standard intuition and the Cox bimodality hypothesis, the relationship between the frequency of strategic voting and the distribution of the vote in the constituency is exactly as specified by the model. Strategic voting increases with the strategic incentive variable derived from the voter uncertainty model and the association is statistically significant.

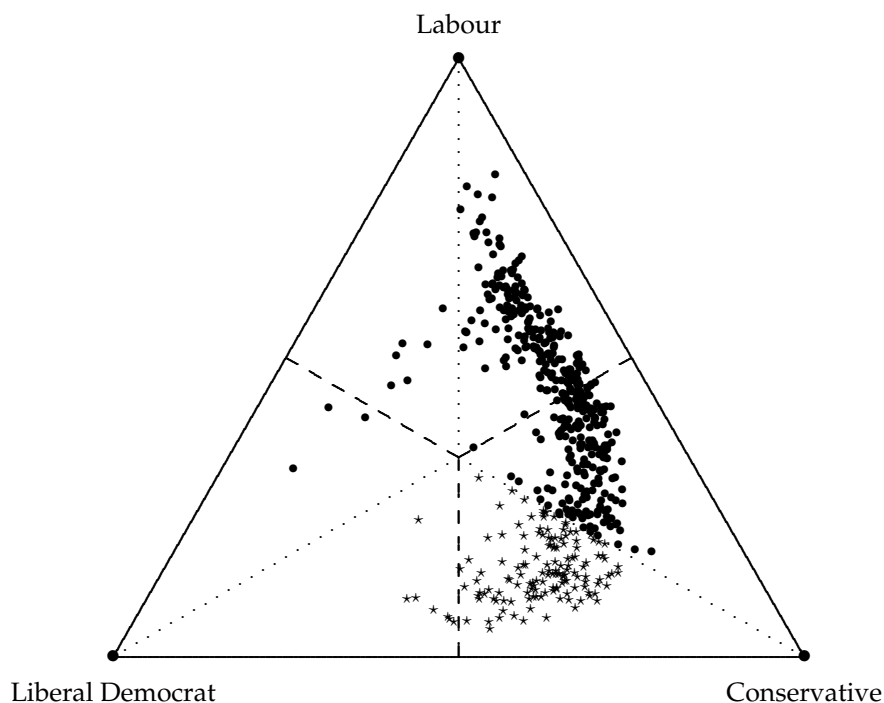
5.3. Cox Bimodality.

Cox (1997) tests for bimodality of the SF ratio (Hypothesis 3) in British parliamentary constituencies, but there are problems with the sample selection procedure, use of marginality, and lack of a null-hypothesis. These will be discussed in turn. First, the test was applied to General Election results between 1983 and 1992 in which Labour came third (e.g. Figure 2). In the conclusion to the test Cox (1997, pp. 88–89) remarks:

“Although the evidence just discussed does indicate that there is strategic voting in *some* British constituencies, the constituencies chosen for inclusion in the analysis were those in which it would have made sense for voters to consider a tactical vote (the strategy of investigation here is similar to that in Blais and Nadeau, 1996).”

Unfortunately this does not provide a justification for the sample selection procedure. It is not true that the chosen constituencies are “those in which it would have made sense for voters to consider a tactical vote”: there are third party supporters and hence strategic incentives in all constituencies. Cox does not argue that the risk population is greater in the constituencies he has chosen. Indeed it is impossible to do so without turning to survey data because some measure of party support distinct from the share of the vote is required. For these reasons the choice

of constituencies seems arbitrary and therefore unacceptable. To find a non-random subset of a population that fits a hypothesis is not to find evidence for the hypothesis.



Constituencies are indicated by either a bullet point “•” or star “*,” where the latter indicate constituencies surviving Cox’s (1997) sample selection procedure. Such points are a weighted average of the three extreme points (which indicate 100% vote for the party labeled). Weights corresponding to the relative vote shares of the major parties. If the Cox bimodality hypothesis is correct, then constituencies should be clustered around the outside of the plot or along the dotted lines. By inspection, they are not. $N=527$.

FIGURE 2. Simplex Plot for English Constituencies in 1992

The second problem is that Cox uses the marginality hypothesis when it is not a prediction derived from his model. Rather, he justifies it according to the standard intuition: strategic voting is greater when the margin of victory is small because the absolute probability of influencing the result is greater. Thus Cox postulates that the tendency to bimodality in the SF ratio will increase as the margin decreases. Histograms with different restrictions on the margin of victory show that the SF ratio becomes more clustered at one and zero when the margin of victory is smaller (Cox 1997, pp.87-8). More precisely, the bimodality hypothesis only seems to hold for marginal constituencies.⁹ So, in addition to the requirement that Labour came third, the margin of victory also

⁹Unfortunately, the histograms are all on different scales. Whilst some aspects of the graphs do clearly distinguish the histograms, others are misleading. For example, Cox (1997) rightly notes that the shape of the distributions in his Figures 4.1 and 4.2 are clearly different, but he draws too strong a conclusion from his comparison of Figures 4.1 and 4.3 (pp. 87-8).

acts as a sample selection criterion which is equally problematic. Since the bimodality hypothesis is generated using formal theory which rejects the marginality hypothesis, it is inconsistent to combine formal and informal theory in this way.

The third problem with Cox's analysis is the absence of a null hypothesis, which is important in order to know what would constitute evidence against the bimodality hypothesis. Cox (1994) argues that there is no apparent reason why the SF ratio should have a bimodal distribution if his theory is not true. While he may be right, the reader should not accept the burden of proof.¹⁰ A related problem is that *any* prior distribution for the SF ratio must vary with the margin of victory. So it is unreasonable to compare the histograms Cox presents as if they should *a priori* be similar; they ought to be different, but it is not clear how different and in what way different.

Whilst the sample selection and use of marginality issues can easily be corrected, the lack of null-hypothesis has more radical implications for the test of the Cox bimodality hypothesis. Two conditions need to be satisfied. Firstly, the SF ratio should have a bimodal distribution. Secondly, if a bimodal distribution is actually the result of strategic voting, then strategic voting should be higher in those constituencies in which the SF ratio is close to zero and lower in those where the SF ratio is near one. More generally, the level of strategic voting should decline as the SF ratio increases.

Addressing the second condition first, Table 3 shows the level of strategic voting at the 1987, 1992 and 1997 British general elections among voters in England only according to different levels of the SF ratio. There was nearly no association between strategic voting and the SF ratio in 1987. For 1992 there appears to be a negative association as hypothesized, but the pattern is not monotonic and logistic regression shows the relationship is not statistically significant. Finally, in 1997 the association is in the opposite direction to that predicted. So even if the aggregate distribution of the SF ratio is bimodal, the survey evidence suggests that it was not generated by strategic voting. Whether or not the SF ratio has the required distribution remains to be shown.¹¹

The observed SF ratios in English constituencies for 1992 are presented in a histogram in Figure 3. The distributions in 1987 and 1997 are very similar but not presented here, since it is only in 1992 that the pattern of strategic voting in the survey data approaches that prescribed by the Cox

¹⁰The specification of a null hypothesis requires a prediction of what constituency results would be if there were no strategic voting. With available survey data, it is impossible to adequately estimate "sincere" results in any particular constituency.

¹¹Strategic votes away from parties that actually came first or second are still included in Table 3. These strategic voters did not affect the SF ratio in the prescribed manner and some may have been confounding the tendency to a Duvergerian outcome. However, if all the strategic voters outside the risk population are excluded the conclusion is still the same.

	1987		1992		1997	
	%	N	%	N	%	N
0 to 0.25	6.0	(156)	10.3	(240)	4.1	(76)
0.25 to 0.5	4.7	(928)	7.6	(1102)	8.5	(849)
0.5 to 0.75	5.1	(985)	8.5	(466)	8.6	(545)
0.75 to 1	5.0	(585)	3.5	(202)	9.5	(333)
Average	5.0	(2653)	7.7	(2010)	8.5	(1802)

Notes: Source is the British Election Studies 1987, 1992 and 1997. Base is voters in England only. The SF ratio for an individual is the SF ratio in their constituency. The first column gives four categories of SF ratio. Remaining columns give the percentage of strategic voters, as well as the total number of voters in each category.

TABLE 3. Percent Strategic by the SF Ratio

theory. According to the Cox hypothesis this histogram of the SF ratio should be bimodal, but it is clearly single-peaked. So there is no evidence that constituencies tend towards two distinct types of equilibria. The conclusion is, therefore, against Cox’s bimodality hypothesis.

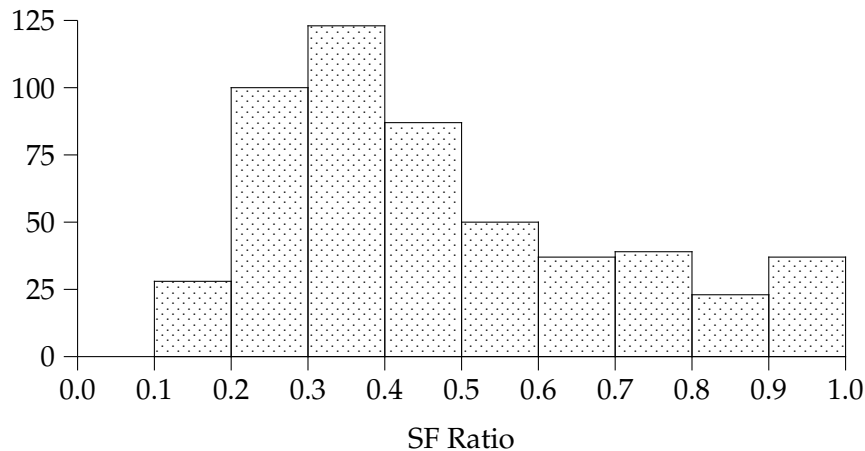


FIGURE 3. Histogram of the SF Ratio for English constituencies in 1992

6. CONCLUSION

We have presented a new approach to strategic voting which incorporates voter uncertainty and tested it together with two alternatives using British Election Study data. Standard intuition supposes that the most likely person to vote strategically will be indifferent between their first and second choice parties, strongly prefer their second to their third choice, and live in a marginal constituency where their preferred party is a long way behind the second placed party. These ideas are largely right but crucially, strategic voting does not increase with marginality after controlling for distance from contention. Cox (1994) argues that in any constituency there will either be nearly

no strategic voting, because the second and third placed parties are close, or the third placed party will be almost completely strategically deserted. In fact, constituencies do not cluster in either of these two categories. Moreover, the pattern of strategic voting in the BES is not one that would push results in the direction supposed by Cox.

The main problem with the Cox-Palfrey approach is that voters are certain of the support for the parties. This assumption drives the extreme predictions of full coordination, or perfectly balanced ties for second place. If voters are allowed to be uncertain the equilibria are not so extreme. This paper presents a model which incorporates voter uncertainty and generates comparative statics that are consistent with the pattern of strategic voting in England in 1987, 1992 and 1997. To this extent the voter uncertainty model is the best account of strategic voting in England yet. Important questions remain, however.

First, the formal model described here is decision-theoretic. Whereas this results in some strategic voting, we might expect game-theoretic considerations to expand this effect. The “bandwagon” logic of strategic voting suggests that the loss of support for trailing candidates will increase strategic incentives, causing a positive feedback loop that leads eventually to a fully Duvergerian conclusion. Related work (Myatt 2002), however, shows that this logic is seriously flawed. When voters are uncertain, they become rather more cautious in their behavior. If they expect others to readily switch their vote in response to perceptions of the constituency situation, then this caution increases. Intuitively, they become concerned that any strategic voting bandwagon may well be rolling in the opposite direction to that indicated by their own information sources. In fact, a formal game-theoretic model shows that strategic voting exhibits negative feedback, and that partial strategic voting and hence multi-candidate support arises as the uniquely stable equilibrium. Furthermore, the comparative statics from such a model support those described here — after controlling for the distance from contention, strategic voting is increasing in the size of the winning margin.

Second, the formal theory is sufficiently complex as to raise important methodological questions. Does it matter that voters do not make the calculations prescribed by the model? It seems odd to think that some voters are consciously voting strategically, but make a sub-conscious strategic decision according to a mechanism that has not, until now, been elucidated by anyone. But, it is even more odd to think that people have been making strategic decisions according to some other mechanism, however intuitively appealing, when there is no evidence for it. The basic components of the model are easily understood and difficult to disagree with: voters can only affect the result when they are pivotal and they are always uncertain of the result. The model

shows how these features of the voting context affect the optimal vote choice for an instrumentally rational voter. Although the predictions are counter intuitive it should not be surprising to find they work, because the premises and the framework of the model are essentially correct.

One general lesson from this paper is that formal theory can be useful for understanding political behavior when intuition fails. Previously, authors such as Green & Shapiro (1994) have claimed that rational choice theorists merely formalize the intuitive and show us what we already knew to be true. Here we have a case in which formal theory contradicts standard intuition and better explains the empirical observations. Another lesson from this paper is that the success of formal rational choice theorizing depends crucially on the components of the model. The failure of the Cox model is due to an unrealistic assumption of voter certainty. Everything is uncertain and uncertainty is everything when it comes to understanding strategic voting incentives.

APPENDIX A. OMITTED PROOFS — LIMITING BEHAVIOR

Lemma 2 and its corollary are used in the proof of Proposition 1.

Lemma 2. *For any $\gamma \in \Delta$:*

$$\lim_{n \rightarrow \infty} \frac{\int_{\Delta} f(p) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta} f(\gamma) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp} = 1 \quad (9)$$

Proof. The function $\prod_{j=1}^m p_j^{\gamma_j}$ is strictly quasi-concave over Δ , with a unique and strict global maximum at $p = \gamma$. For arbitrarily small $\epsilon < \prod_{j=1}^m \gamma_j^{\gamma_j}$, define:

$$\Delta_{\epsilon} = \left\{ p \in \Delta \mid \prod_{j=1}^m p_j^{\gamma_j} \geq \prod_{j=1}^m \gamma_j^{\gamma_j} - \epsilon \right\}$$

This is a compact and convex subset of Δ . We first construct a lower bound to the left hand side of Equation 9. Begin by noting that:

$$\int_{\Delta} f(p) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp \geq \underline{f}_{\epsilon} \int_{\Delta_{\epsilon}} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp \quad \text{where} \quad \underline{f}_{\epsilon} = \min_{p \in \Delta_{\epsilon}} f(p)$$

The minimum is well defined, from compactness of Δ_{ϵ} and continuity of $f(p)$. It follows:

$$\begin{aligned} \frac{\int_{\Delta} f(p) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta} f(\gamma) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp} &\geq \frac{\underline{f}_{\epsilon}}{f(\gamma)} \frac{\int_{\Delta_{\epsilon}} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta_{\epsilon}} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp + \int_{\Delta/\Delta_{\epsilon}} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp} \\ &= \frac{\underline{f}_{\epsilon}}{f(\gamma)} \frac{1}{1 + \left[\int_{\Delta/\Delta_{\epsilon}} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp \right] / \left[\int_{\Delta_{\epsilon}} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp \right]} \end{aligned}$$

The next step is to show that the ratio of integrals in the denominator of this expression tends to zero for large n . To see this:

$$\frac{\int_{\Delta/\Delta_\epsilon} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta_\epsilon} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp} \leq \frac{\int_{\Delta/\Delta_\epsilon} \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta_\epsilon} \left[\prod_{j=1}^m \gamma_j^{\gamma_j} - \epsilon \right]^n dp} = \frac{1}{\int_{\Delta_\epsilon} dp} \int_{\Delta/\Delta_\epsilon} \left[\frac{\prod_{j=1}^m p_j^{\gamma_j}}{\prod_{j=1}^m \gamma_j^{\gamma_j} - \epsilon} \right]^n dp$$

Of course, for $p \notin \Delta_\epsilon$ it must be that $\prod_{j=1}^m p_j^{\gamma_j} < \prod_{j=1}^m \gamma_j^{\gamma_j} - \epsilon$. Hence:

$$\int_{\Delta/\Delta_\epsilon} \left[\frac{\prod_{j=1}^m p_j^{\gamma_j}}{\prod_{j=1}^m \gamma_j^{\gamma_j} - \epsilon} \right]^n dp \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \frac{\int_{\Delta} f(p) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta} f(\gamma) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp} \geq \frac{f_\epsilon}{f(\gamma)}$$

It is clear that $\gamma \in \Delta_\epsilon$. For sufficiently small ϵ , Δ_ϵ is an arbitrarily small neighborhood of γ , since γ is the unique and strict global maximizer of the continuous function $\prod_{j=1}^m p_j^{\gamma_j}$. Thus, for a sufficiently small ϵ , f_ϵ may be drawn arbitrarily close to $f(\gamma)$, implying:

$$\lim_{n \rightarrow \infty} \frac{\int_{\Delta} f(p) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp}{\int_{\Delta} f(\gamma) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp} \geq 1$$

This proof constructed a lower bound to the required expression. A symmetric procedure yields an upper bound with the same properties, and the result obtains. \square

Corollary 1. For larger constituencies, outcome probabilities satisfy:

$$\int_{\Delta} f(p) \prod_{j=1}^m p_j^{x_j} dp \underset{n \rightarrow \infty}{\approx} f\left(\frac{x}{n}\right) \int_{\Delta} \prod_{j=1}^m p_j^{x_j} dp$$

Proof. From the compactness of Δ , uniform convergence means that:

$$\int_{\Delta} f(p) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp \underset{n \rightarrow \infty}{\approx} \int_{\Delta} f(\gamma) \left[\prod_{j=1}^m p_j^{\gamma_j} \right]^n dp$$

Defining $\gamma_j = x_j/n$, the result follows. \square

Proof of Proposition 1. From an application of Corollary 1 to the expression for $\Pr[x]$:

$$\Pr[x] = \frac{\Gamma(n+1)}{\prod_{j=1}^m \Gamma(x_j+1)} \int_{\Delta} f(p) \prod_{j=1}^m p_j^{x_j} dp \underset{n \rightarrow \infty}{\approx} \frac{\Gamma(n+1)}{\prod_{j=1}^m \Gamma(x_j+1)} \int_{\Delta} f\left(\frac{x}{n}\right) \prod_{j=1}^m p_j^{x_j} dp$$

where $\Gamma(\cdot)$ is the Gamma function. Multiplying and dividing by $\Gamma(n+m)$ obtain:

$$\Pr[x] \underset{n \rightarrow \infty}{\approx} f\left(\frac{x}{n}\right) \frac{\Gamma(n+1)}{\Gamma(n+m)} \int_{\Delta} \frac{\Gamma(n+m)}{\prod_{j=1}^m \Gamma(x_j+1)} \prod_{j=1}^m p_j^{x_j} dp = f\left(\frac{x}{n}\right) \frac{\Gamma(n+1)}{\Gamma(n+m)}$$

The last step follows from recognition of the last integrand as a Dirichlet density. Both n and m are integers. Recalling that $\Gamma(n+1) = n!$ and $\Gamma(n+m) = (n+m-1)!$:

$$\frac{\Gamma(n+1)}{\Gamma(n+m)} = \frac{1}{\prod_{k=1}^{m-1} (n+k)} \underset{n \rightarrow \infty}{\approx} \frac{1}{n^{m-1}}$$

From this the result follows. □

Proof of Lemma 1. Focusing on the case of a two-way tie between candidates 1 and 2:

$$\Pr[x_1 = x_2 > x_3] = \sum_{\frac{n}{3} < i \leq \frac{n}{2}} \Pr[i, i, n-2i]$$

where $\Pr[i, i, n-2i] = \Pr[x_1 = x_2 = i \text{ and } x_3 = n-2i]$. Multiply through by n :

$$n \Pr[x_1 = x_2 > x_3] = \frac{1}{n} \sum_{\frac{n}{3} < i \leq \frac{n}{2}} n^2 \Pr[i, i, n-2i] \underset{n \rightarrow \infty}{\approx} \frac{1}{n} \sum_{\frac{n}{3} < i \leq \frac{n}{2}} f\left(\frac{i}{n}, \frac{i}{n}, \frac{n-2i}{n}\right)$$

The sum on the right hand side has approximately $n/6$ elements. Multiplying and dividing by 6, this summation defines a Riemann integral over the set $[1/3, 1/2]$. Indeed:

$$n \Pr[x_1 = x_2 > x_3] \underset{n \rightarrow \infty}{\approx} \frac{1}{6} \frac{1}{n/6} \sum_{\frac{n}{3} < i \leq \frac{n}{2}} f\left(\frac{i}{n}, \frac{i}{n}, \frac{n-2i}{n}\right) \rightarrow \frac{1}{6} \int_{1/3}^{1/2} f(z, z, 1-2z) dz$$

This proves the part of Lemma 1. Three-way ties follow directly from Proposition 1. □

APPENDIX B. OMITTED PROOFS — PROPERTIES OF DIRICHLET BELIEFS

The basic properties $g(z, \pi)$ from Equation 7 are used in the proof of Proposition 3.

Lemma 3. For $1/3 < z < 1/2$, the function $g(z, \pi)$ is strictly decreasing in π and satisfies $0 < g(z, \pi) < 1$. At the endpoints of the interval $g(1/3, \pi) = 1/3$ and $g(1/2, \pi) = 0$.

Proof. To show that $g(z, \pi)$ is decreasing in π , write:

$$g(z, \pi) = z \left[\frac{1-2z}{z} \right]^\pi \Rightarrow \frac{\partial g(z, \pi)}{\partial \pi} = g(z, \pi) \log \left[\frac{1-2z}{z} \right] < 0 \quad \text{for} \quad \frac{1}{3} < z < \frac{1}{2}$$

The remaining properties hold by inspection. □

Proof of Proposition 3. From Proposition 2 and Equation 6:

$$\begin{aligned} p_{12} &= \int_{1/3}^{1/2} f(z, z, 1 - 2z) dz \propto \int_{1/3}^{1/2} [z^{\pi_1} z^{\pi_2} (1 - 2z)^{\pi_3}]^s dz \\ &= \int_{1/3}^{1/2} [z^{1-\pi_3} (1 - 2z)^{\pi_3}]^s dz = \int_{1/3}^{1/2} g(z, \pi_3)^s dz \end{aligned}$$

Similar operations for p_{13} and p_{23} yield the first part of Proposition 3. To obtain the explicit expression, for $j \in \{1, 2, 3\}$, make the change of variable $w = 1 - 2z$ to obtain:

$$\begin{aligned} \int_{1/3}^{1/2} [z^{1-\pi_j} (1 - 2z)^{\pi_j}]^s dz &= \frac{1}{2^{(1-\pi_j)s+1}} \int_0^{1/3} [w^{\pi_j} (1 - w)^{1-\pi_j}]^s dw \\ &= \frac{2^{\pi_j s} B_{1/3}(\pi_j s + 1, (1 - \pi_j)s + 1)}{2^{s+1}} \end{aligned}$$

where the last equality follows from recognizing the integral expression as the incomplete Beta function evaluated at $1/3$ with the specified parameters. Form the strategic incentive λ , and cancel appropriate terms to obtain:

$$\begin{aligned} \lambda &= \log \frac{2^{1+\pi_3 s} B_{1/3}(\pi_3 s + 1, (1 - \pi_3)s + 1) + 2^{\pi_1 s} B_{1/3}(\pi_1 s + 1, (1 - \pi_1)s + 1)}{2^{1+\pi_3 s} B_{1/3}(\pi_3 s + 1, (1 - \pi_3)s + 1) + 2^{\pi_2 s} B_{1/3}(\pi_2 s + 1, (1 - \pi_2)s + 1)} \\ &= \log \frac{B_{1/3}(\pi_3 s + 1, (1 - \pi_3)s + 1) + 2^{(\pi_1 - \pi_3)s-1} B_{1/3}(\pi_1 s + 1, (1 - \pi_1)s + 1)}{B_{1/3}(\pi_3 s + 1, (1 - \pi_3)s + 1) + 2^{(\pi_2 - \pi_3)s-1} B_{1/3}(\pi_2 s + 1, (1 - \pi_2)s + 1)} \end{aligned}$$

which is the desired expression. Turning to comparative statics, from the definition of λ :

$$\lambda > 0 \Leftrightarrow \frac{\int_{1/3}^{1/2} [2g(z, \pi_3)^s + g(z, \pi_1)^s] dz}{\int_{1/3}^{1/2} [2g(z, \pi_3)^s + g(z, \pi_2)^s] dz} > 1 \Leftrightarrow \int_{1/3}^{1/2} g(z, \pi_1)^s dz > \int_{1/3}^{1/2} g(z, \pi_2)^s dz$$

From Lemma 3, if $\pi_2 > \pi_1$ then $g(z, \pi_1) > g(z, \pi_2)$ for $z \in [1/3, 1/2]$. The reverse argument holds for $\pi_1 < \pi_2$, and $\lambda = 0$ when $\pi_1 = \pi_2$. \square

The proof of Proposition 4 employs further properties of $g(z, \pi)$ from Equation 7.

Lemma 4. *The expression $\log \left[\int_{1/3}^{1/2} g(z, \pi)^s dz \right]$ is convex in π .*

Proof. Consider the first derivative of the expression:

$$\frac{\partial \log \left[\int_{1/3}^{1/2} g(z, \pi)^s dz \right]}{\partial \pi} = \frac{1}{\int_{1/3}^{1/2} g(z, \pi)^s dz} \int_{1/3}^{1/2} s g(y, \pi)^{s-1} \frac{\partial g(y, \pi)}{\partial \pi} dy$$

where we have used to different variables of integration for each separate integral in order to avoid confusion. Turning to the function $g(y, \pi)$:

$$g(y, \pi) = y^{1-\pi}(1-2y)^\pi \quad \Rightarrow \quad \frac{\partial g(y, \pi)}{\partial \pi} = g(y, \pi) \log \left[\frac{1-2y}{y} \right]$$

which on substitution and multiplication through by -1 yields:

$$-\frac{\partial \log \left[\int_{1/3}^{1/2} g(z, \pi)^s dz \right]}{\partial \pi} = s \int_{1/3}^{1/2} \log \left[\frac{y}{1-2y} \right] \left[\frac{g(y, \pi)^s}{\int_{1/3}^{1/2} g(z, \pi)^s dz} \right] dy \quad (10)$$

The multiplication by -1 ensures that we are dealing with a positive expression, since $y > 1-2y$ on the required interval. We wish to show that that this expression is decreasing in π . Now, observe that $\log[y/1-2y]$ is strictly increasing in y . Also observe that:

$$\frac{g(y, \pi)^s}{\int_{1/3}^{1/2} g(z, \pi)^s dz} \quad (11)$$

is a well defined density function — it is positive valued, and integrates to 1 over the interval $[1/3, 1/2]$. It follows that the expression in Equation 10 is the expectation of an increasing function. It is thus sufficient to show that a reduction in π induces a first order stochastically dominant shift upwards in the density described by Equation 11. To demonstrate this, I must show that, for any y where $1/3 < y < 1/2$:

$$\begin{aligned} \pi_j < \pi_k &\Rightarrow \frac{\int_{1/3}^y g(w, \pi_j)^s dw}{\int_{1/3}^{1/2} g(z, \pi_j)^s dz} \leq \frac{\int_{1/3}^y g(w, \pi_k)^s dw}{\int_{1/3}^{1/2} g(z, \pi_k)^s dz} \\ &\Leftrightarrow \frac{\int_{1/3}^y g(w, \pi_j)^s dw}{\int_{1/3}^y g(w, \pi_k)^s dw} \leq \frac{\int_{1/3}^{1/2} g(z, \pi_j)^s dz}{\int_{1/3}^{1/2} g(z, \pi_k)^s dz} \end{aligned}$$

To demonstrate that this inequality holds it is sufficient to show that the left hand side term is maximized at $y = 1/2$. To show this, it is sufficient to show that its derivative is positive on the interval $1/3 < y < 1/2$. To show this requires:

$$\begin{aligned} \frac{\partial}{\partial y} \log \left[\frac{\int_{1/3}^y g(w, \pi_j)^s dw}{\int_{1/3}^y g(w, \pi_k)^s dw} \right] \geq 0 &\Leftrightarrow \frac{g(y, \pi_j)^s}{\int_{1/3}^y g(w, \pi_j)^s dw} \geq \frac{g(y, \pi_k)^s}{\int_{1/3}^y g(w, \pi_k)^s dw} \\ &\Leftrightarrow \frac{\int_{1/3}^y g(w, \pi_j)^s dw}{g(y, \pi_j)^s} \leq \frac{\int_{1/3}^y g(w, \pi_k)^s dw}{g(y, \pi_k)^s} \end{aligned}$$

This last inequality holds if, for all w and y satisfying $1/3 < w \leq y < 1/2$:

$$\frac{g(w, \pi_j)}{g(y, \pi_j)} \leq \frac{g(w, \pi_k)}{g(y, \pi_k)} \quad \Leftrightarrow \quad \left[\frac{y(1-2w)}{w(1-2y)} \right]^{\pi_j} \leq \left[\frac{y(1-2w)}{w(1-2y)} \right]^{\pi_k}$$

This is true, since $\pi_j < \pi_k$ and $y(1 - 2w) \geq w(1 - 2y)$ for $1/3 < w \leq y < 1/2$. Assembling calculations so far, I have demonstrated that the right hand side of Equation 10 is decreasing in π . It follows that:

$$\frac{\partial^2 \log \left[\int_{1/3}^{1/2} g(z, \pi)^s dz \right]}{\partial \pi^2} \geq 0$$

which is the required result. \square

Before giving the proof of Proposition 4, we review the formal definitions of c and d , and invert to obtain π_j in terms of these parameters. Consider a configuration where the preferred candidate trails in third place, so either $\pi_2 > \pi_3 > \pi_1$ or $\pi_3 > \pi_2 > \pi_1$. As noted in the text, the distance from contention may be defined as $d = \min\{\pi_2, \pi_3\} - \pi_1$ and the winning margin for the most disliked candidate may be defined as $c = \pi_3 - \pi_2$, so that $|c|$ is absolute size of the winning margin. Using these definitions, and the requirement that $\pi_1 + \pi_2 + \pi_3 = 1$, π can be defined in terms of c and d . Solving the relevant equations linearly yields the solutions:

	$c > 0$	$c < 0$
$\pi_1 = \frac{1}{3} \times$	$1 - 2d - c$	$1 - 2d + c$
$\pi_2 = \frac{1}{3} \times$	$1 + d - c$	$1 + d - 2c$
$\pi_3 = \frac{1}{3} \times$	$1 + d + 2c$	$1 + d + c$

(12)

Notice that the behavior of π (and hence the strategic incentive λ) in response to the winning margin c changes at $c = 0$, since at this point the ordering of the candidates changes. We are now ready to prove Proposition 4.

Proof of Proposition 4. First, notice that the distance from contention effect is as predicted by intuition. Inspecting the solutions for π in terms of c and d from Equation 12 above, and fixing the winning margin c , an increase in d lowers π_1 and simultaneously increases both π_2 and π_3 . It follows from Proposition 3 that λ rises. The major task of the proof, therefore, is to assess the effect of c when d is fixed.

Consider first the case where $\pi_3 > \pi_2 > \pi_1$, so that $c > 0$. Taking the solutions for π in terms of c and d from Equation 12:

$$\frac{\partial \pi_1}{\partial c} = -\frac{1}{3}, \quad \frac{\partial \pi_2}{\partial c} = -\frac{1}{3} \quad \text{and} \quad \frac{\partial \pi_3}{\partial c} = \frac{2}{3} \quad (13)$$

Next consider the strategic incentive λ . Using Proposition 2:

$$\lambda_\infty = \log \frac{2p_{12} + p_{23}}{2p_{12} + p_{13}} = \log \frac{2(p_{12}/p_{13}) + (p_{23}/p_{13})}{2(p_{12}/p_{13}) + 1}$$

Since $\pi_2 > \pi_1$, the strategic incentive is positive. Equivalently, $p_{23} > p_{13}$. It follows that λ is decreasing in (p_{12}/p_{13}) . Using the results of Proposition 3:

$$\frac{p_{12}}{p_{13}} = \frac{\int_{1/3}^{1/2} g(z, \pi_3)^s dz}{\int_{1/3}^{1/2} g(z, \pi_2)^s dz}$$

From Equation 13 above, an increase in c reduces π_2 and increases π_3 . Using the results of Lemma 3, this increases $g(z, \pi_2)$ and reduces $g(z, \pi_3)$. Combining, it follows that (p_{12}/p_{13}) is decreasing in c . This means that, to demonstrate the claim of the Proposition, it is sufficient to show that (p_{23}/p_{13}) is increasing in c . To do this, differentiate to obtain:

$$\begin{aligned} \frac{\partial \log(p_{23}/p_{13})}{\partial c} &= \frac{\partial \log \int_{1/3}^{1/2} g(z, \pi_1)^s dz}{\partial \pi_1} \frac{\partial \pi_1}{\partial c} - \frac{\partial \log \int_{1/3}^{1/2} g(z, \pi_2)^s dz}{\partial \pi_2} \frac{\partial \pi_2}{\partial c} \\ &= \frac{1}{3} \left\{ \frac{\partial \log \int_{1/3}^{1/2} g(z, \pi_2)^s dz}{\partial \pi_2} - \frac{\partial \log \int_{1/3}^{1/2} g(z, \pi_1)^s dz}{\partial \pi_1} \right\} > 0 \end{aligned}$$

The last inequality follows from Lemma 4 and the fact that $\pi_2 > \pi_1$ by assumption. Combining, it follows that λ is increasing in c for $c > 0$. Turning to the case of $c < 0$, a similar proof yields the desired result. \square

The proof of Proposition 5 is based upon the following lemma.

Lemma 5. Define $g^*(\pi)$ as the maximum of $g(z, \pi)$ on $1/3 \leq \pi \leq 1/2$.

$$\lim_{s \rightarrow \infty} \left\{ \frac{\log \left[\int_{1/3}^{1/2} g_j(z)^s dz \right]}{s \log g^*(\pi)} \right\} = 1$$

Proof. Both numerator and denominator tend to $-\infty$. Apply l'Hôpital's rule to obtain:

$$\lim_{s \rightarrow \infty} \left\{ \frac{\log \left[\int_{1/3}^{1/2} g_j(z)^s dz \right]}{s \log g^*(\pi)} \right\} = \frac{1}{\log g^*(\pi)} \times \lim_{s \rightarrow \infty} \left\{ \int_{1/3}^{1/2} \log g(z, \pi) \left[\frac{g(z, \pi)^s}{\int_{1/3}^{1/2} g(y, \pi)^s dy} \right] dz \right\}$$

The second term in the integrand defines a density function. As $s \rightarrow \infty$, it focuses increasing weight around the maximum of $g(z, \pi)$, and hence the integral tends to $g^*(z)$. A full formal proof of this follows an identical structure to the proof of Lemma 2. \square

Proof of Proposition 5: Write λ as:

$$\lambda = \log \frac{2(p_{12}/p_{23}) + 1}{2(p_{12}/p_{23}) + (p_{13}/p_{23})} \quad (14)$$

Take, for example the ratio (p_{13}/p_{23}) . Apply Lemma 5 to obtain:

$$\log \frac{p_{13}}{p_{23}} = \log \frac{\int_{1/3}^{1/2} g(z, \pi_2)^s dz}{\int_{1/3}^{1/2} g(z, \pi_1)^s dz} \longrightarrow s [\log g^*(\pi_2) - \log g^*(\pi_1)] \quad \text{as } s \rightarrow \infty$$

When $\pi_1 < \min\{\pi_2, \pi_3\}$ it is straightforward to observe that $\pi_1 < 1/3$ and show that $g^*(\pi_1) > \max\{g^*(\pi_2), g^*(\pi_3)\}$. It follows that $(p_{13}/p_{23}) \rightarrow 0$ as $s \rightarrow \infty$. Similar operations on the remaining elements of Equation 14 demonstrate that $\lambda \rightarrow \infty$. A similar proof applies to the case of $\pi_3 < \pi_1 < \pi_2$. \square

APPENDIX C. NUMERICAL CALCULATIONS

We briefly describe the implementation of our strategic incentive variable. Using the Dirichlet density and applying Proposition 3 the strategic incentive solves (Equation 8):

$$\lambda_\infty = \log \frac{B_{1/3}(\pi_3 s + 1, (1 - \pi_3)s + 1) + 2^{(\pi_1 - \pi_3)s - 1} B_{1/3}(\pi_1 s + 1, (1 - \pi_1)s + 1)}{B_{1/3}(\pi_3 s + 1, (1 - \pi_3)s + 1) + 2^{(\pi_2 - \pi_3)s - 1} B_{1/3}(\pi_2 s + 1, (1 - \pi_2)s + 1)}$$

where $B_{1/3}(a, b)$ is the incomplete Beta function evaluated at $1/3$. This satisfies:

$$B_{1/3}(a, b) = \int_0^{1/3} z^{a-1} (1-z)^{b-1} dz = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \times \underbrace{\int_0^{1/3} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} z^{a-1} (1-z)^{b-1} dz}_{\text{Regularized Incomplete Beta Function}}$$

where $\Gamma(\alpha)$ is the Gamma function. Of course, the regularized incomplete Beta function is the cumulative distribution function of the Beta distribution. Microsoft Excel implements this using the BETADIST function. Furthermore, the Gamma function is implemented via GAMMALN, satisfying $\text{GAMMALN}(\alpha) = \log \Gamma(\alpha)$. Hence, using Excel notation:

$$B_{1/3}(a, b) = \text{EXP}(\text{GAMMALN}(a) + \text{GAMMALN}(b) - \text{GAMMALN}(a + b)) * \text{BETADIST}(1/3, a, b)$$

Manipulating the Γ components assists with numerical accuracy when using Excel. An alternative to Excel is a full matrix language such as MATLAB. MATLAB once again implements the regularised incomplete function using the function BETAINC. In addition, MATLAB also defines the Beta function $\text{BETA}(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$. Hence:

$$B_{1/3}(a, b) = \text{BETA}(1/3, a, b) * \text{BETAINC}(1/3, a, b)$$

in MATLAB notation. These procedures may be used to generate the strategic incentive variable used in the empirical analysis of Section 5.

REFERENCES

- Abramson, Paul R., John H. Aldrich, Phil Paolino & David W. Rohde. 1992. "'Sophisticated' Voting in the 1988 Presidential Primaries." *American Political Science Review* 86(1):55–69.
- Blais, Andre & R. Nadeau. 1996. "Measuring Strategic Voting: A Two-Step Procedure." *Electoral Studies* 15:39–52.
- Cain, Bruce E. 1978. "Strategic Voting in Britain." *American Journal of Political Science* 22(3):639–655.
- Cox, Gary W. 1994. "Strategic Voting Equilibria under the Single Nontransferable Vote." *American Political Science Review* 88(3):608–621.
- Cox, Gary W. 1997. *Making Votes Count*. Cambridge: Cambridge University Press.
- Curtice, John & Michael Steed. 1988. Appendix 2: Analysis. In *The British General Election of 1987*, ed. David Butler & Dennis Kavanagh. Basingstoke: Macmillan.
- Curtice, John & Michael Steed. 1992. Appendix 2: The Results Analysed. In *The British General Election of 1992*, ed. David Butler & Dennis Kavanagh. Basingstoke: Macmillan.
- Curtice, John & Michael Steed. 1997. Appendix 2: The Results Analysed. In *The British General Election of 1997*, ed. David Butler & Dennis Kavanagh. Basingstoke: Macmillan.
- Duverger, Maurice. 1954. *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.
- Evans, Geoffrey. 1994. Tactical Voting and Labour's prospects. In *Labour's Last Chance? The 1992 Election and Beyond*, ed. Anthony Heath, Roger Jowell & John Curtice. Aldershot: Dartmouth.
- Evans, Geoffrey & Anthony Heath. 1993. "A Tactical Error in the Analysis of Tactical Voting: A Response to Niemi, Whitten and Franklin." *British Journal of Political Science* 23:131–137.
- Farquharson, Robin. 1969. *Theory of Voting*. New Haven: Yale University Press.
- Fey, Mark. 1997. "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting." *American Political Science Review* 91(1):135–147.
- Fieldhouse, E. A., C. J. Pattie & R. J. Johnston. 1996. "Tactical Voting and Party Constituency Campaigning at the 1992 General Election in England." *British Journal of Political Science* 26:403–439.
- Franklin, Mark, Richard G. Niemi & Guy Whitten. 1994. "The Two Faces of Tactical Voting." *British Journal of Political Science* 24:549–557.
- Green, Donald P. & Ian Shapiro. 1994. *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, CT: Yale University Press.
- Heath, Anthony & Geoffrey Evans. 1994. "Tactical Voting: Concepts, Measurements and Findings." *British Journal of Political Science* 24:557–561.
- Heath, Anthony, John Curtice, Roger Jowell, Geoff Evans, Julia Field & Sharon Witherspoon. 1991. *Understanding Political Change: The British Voter 1964–1987*. Oxford: Pergamon Press.
- Heath, Anthony, Roger Jowell & John Curtice. 1995. *British Election Study: Cross-Section, 1987 [Computer File] (Study #6452)*. London, England: Social and Community Planning Research/Oxford, England: A. Heath et al., Oxford University [producers], 1989. Colchester,

- England: ESRC Data Archive/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors].
- Heath, Anthony, Roger Jowell, John K. Curtice, J.A. Brand & J.C. Mitchell. 1996. *British General Election Cross-Section Survey, 1992 [Computer File]. (Study #6453) ICPSR Version*. London, England: Social and Community Planning Research [producer], 1992. Colchester, England: ESRC Data Archive/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors].
- Heath, Anthony, Roger Jowell, John K. Curtice & Pippa Norris. 2000. *British General Election Cross-Section Survey, 1997 [Computer File] (Study #2615). ICPSR 2nd Version*. London, England: Social and Community Planning Research [producer], 1998. Colchester, England: The Data Archive/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors].
- Hoffman, Dale T. 1982. "A Model for Strategic Voting." *SIAM Journal of Applied Mathematics* 42:751–761.
- McKelvey, Richard & Peter Ordeshook. 1972. A General Theory of the Calculus of Voting. In *Mathematical Applications in Political Science*, ed. J. F. Herndon & J. L. Bernd. Charlottesville: University Press of Virginia. volume 6.
- Myatt, David P. 2002. "A New Theory of Strategic Voting." University of Oxford Department of Economics Discussion Paper Series no. 93.
- Myerson, Roger B. & Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review* 87(1):102–114.
- Niemi, R., G. Whitten & M. Franklin. 1992. "Constituency Characteristics, Individual Characteristics and Tactical Voting in the 1987 British General Election." *British Journal of Political Science* 22:229–240.
- Niemi, Richard G., Guy Whitten & Mark Franklin. 1993. "People Who Live in Glass Houses: A Response to Evans and Heath's Critique of our Note on Tactical Voting." *British Journal of Political Science* 23:549–563.
- Norris, Pippa, John Curtice, David Sanders, Margaret Scammell & Holli A. Semetko. 1999. *On Message: Communicating the Campaign*. London: Sage.
- Palfrey, Thomas R. 1989. A Mathematical Proof of Duverger's Law. In *Models of Strategic Choice in Politics*, ed. Peter C. Ordeshook. Ann Arbor: University of Michigan Press pp. 69–92.
- Riker, William H. & Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62:25–42.
- Schattschneider, E. E. 1942. *Party Government*. New York: Holt, Reinhardt and Winston.