Strategic Voting in Plurality Rule Elections

Stephen D. Fisher
University of Oxford

David P. Myatt
London Business School

July 10, 2017.

Abstract. Some have suggested that strategic voting should be most likely when a preferred candidate trails behind two narrowly separated competitors. Others have proposed a bi-modality hypothesis: either desertion of all but two candidates, or an exact tie for second place. Our formal theoretical analysis of the incentive to vote strategically in three-candidate elections allows for aggregate uncertainty over the candidates' popularities. We predict multi-candidate support and that, after controlling for the distance from contention of a preferred candidate, strategic voting should be more muted in close-race districts. We examine voters in England from recent General Elections. Our predictions find support in the data, while the informal argument and the bi-modality hypothesis do not. A simple calibration exercise suggests that the precision of information available to voters is moderately low; equivalent to observing a random sample of around 10–12 voting intentions from other members of a voting district. That same exercise suggests that only some voters (approximately 35–40%) are instrumentally motivated.

It is sometimes argued informally that a voter's desire to vote strategically will be strong if her preferred candidate is expected to trail behind other closely matched candidates; that is, if the "distance from contention" (of the preferred candidate) is large and the expected "margin of victory" (between the others) is small. Our formal analysis of three-candidate plurality-rule elections allows for aggregate uncertainty over the candidates' popularities. After controlling for the distance from contention, we show that the incentive to vote strategically should be higher (not lower) in close-race districts. Furthermore, the size of this marginality effect should be relatively weak. We find empirical support from election outcomes in English constituencies. Furthermore, our model explicitly predicts the propensity to vote strategically in different constituency situations and so we are able to calculate the theoretical likelihood of strategic voting across England. A calibration of our model suggests that the pattern and incidence of strategic voting might best be explained by a world in which some but not all voters are instrumentally motivated, and where instrumental voters face substantial uncertainty about their district situations.

Contact. stephen.fisher@sociology.ox.ac.uk (Fisher) and dmyatt@london.edu (Myatt).

Acknowledgements. We thank Jean-Pierre Benoît, Micael Castanheira, Torun Dewan, Andy Eggers, David Firth, Anthony Heath, Libby Hunt, Justin Johnson, Paul Klemperer, Clare Leaver, Iain McLean, Joey McMurray, Adam Meirowitz, Tiago Mendes, Becky Morton, Kevin Roberts, David Sanders, Norman Schofield, Ken Shepsle, Chris Wallace, Stephane Wolton, and Peyton Young for their comments on this and other closely related work. This paper resurrects and incorporates elements of a very old conference and working paper (Fisher and Myatt, 2002a).

The "distance from contention" and "marginality" hypotheses are motivated by the absolute likelihood of pivotal events. However, what should matter is the relative (not absolute) likelihood of such events; rather than ask "will I be pivotal?" a voter should ask "which candidates will tie if I do turn out to be pivotal?" In a close race the probability of a tie between the leading candidates may be higher than in a safe district, but a tie involving a trailing candidate may be more likely too. Informal arguments say little about the relative size of these effects.

In established formal models (Palfrey, 1989; Myerson and Weber, 1993; Cox, 1994) voters' decisions are independent draws from a known distribution, and so (in a large electorate) any tie almost always involves the two leading candidates. This implies that a voter should vote for her favorite amongst these two. However, plurality-rule systems generally exhibit multi-candidate support: beyond the United Kingdom cases documented here, examples include Canada, India, and (formerly) New Zealand (Chhibber and Kollman, 1998; Gaines, 1999; Cox and Schoppa, 2002; Diwakar, 2007). To obtain multi-candidate support, Cox (1994) constructed an equilibrium in which a precise fraction of supporters of a leading candidate switch to a trailing candidate to create an exact tie for second place. This requires exact knowledge of the candidates' true popularities, and the topsy-turvy nature of insincere voting means that such equilibria are unstable (Fey, 1997). Nevertheless, such equilibria underpin the bi-modality hypothesis (Cox, 1997) that the "SF ratio" of support for the second to the first loser should be close to zero or to one.

These existing theories do not offer comparative-static predictions, and the data do not support the bi-modality of the SF ratio. These theoretical models also rely on voters' perfect knowledge of their district situation. Cox (1997) commented that a condition "necessary to generate pure local bi-partism 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 . . . " Empirically, voters do have imperfect and different beliefs (Evans and Heath, 1993; Abramson, Aldrich, Paolino, and Rohde, 1992) and so we seek to open up this degree of freedom.

We present a formal theoretical analysis of strategic voting incentives in the presence of aggregate uncertainty about the candidates' popularities. We find that the incentive to vote for one of the two perceived leaders is not overwhelming, and so not everyone votes strategically. One comparative-static prediction is uncontroversial: fixing the expected margin of victory, the strategic-voting incentive grows with the distance from contention of a preferred candidate. However, another prediction overturns the marginality hypothesis: if the distance from contention is fixed then the incentive to vote strategically is lower, not higher, in closer races. We also find that the marginality effect is substantially weaker in size than the distance-from-contention effect.

Our predictions also hold in a game-theoretic setting. We extend the model of Myatt (2007) to incorporate a fraction of non-instrumental voters and to develop further comparative-static results. The marginality effect is again opposite to that of traditional intuition, and it is weaker than the distance-from-contention effect. Furthermore, the willingness of instrumental voters to act strategically is inversely related to the fraction of instrumental voters within the electorate.

The constituencies of England offer a prime example of three-party competition under the simple plurality rule. We investigate comparative-static predictions using data from the 1987, 1992, 1997, 2001, 2005, and 2010 British Election Studies. Those data are inconsistent with the marginality hypothesis of the informal intuition. Instead, the data fit the pattern of strategic incentives generated by our voter uncertainty model: after we control for the distance from contention of a preferred candidate (this is a strong predictor of strategic voting) we find a (much weaker in strength) increase in strategic voting in response to a wider margin of victory.

Our formal model generates a strategic incentive variable which varies with the expectations of the candidates' vote shares and the precision of beliefs. We are able to calculate the distribution of that incentive across English constituencies and so predict the likely incidence of strategic voting. A calibration exercise suggests that the belief precision is moderately low (equivalent to that obtained by sampling around ten independent voting intentions) and that not all voters (a little over one third) are instrumentally motivated. This suggests the use of formal models which incorporate both significant aggregate uncertainty and also a broader range of voter motivations.[3]

A Voting Model with Aggregate Uncertainty

The Election. In a three-candidate plurality election, a voter ("she") contemplates the behavior of n others. Amongst them, x_j is the number of votes cast for candidate $j \in \{1, 2, 3\}$. The three vote counts form the vector $x \in X$, where X is the set of possible outcomes. The candidate receiving the most votes (including the $(n + 1)$th ballot) wins. Ties are broken at random.

[3] Recent experimental and empirical evidence also suggests that aggregate uncertainty (Bouton, Castanheira, and Llorente-Saguer, 2017) and mixed voter motivations (Spenkuch, 2017) are both important.

Beliefs. An IID (independent and identically distributed) specification for a voter's beliefs is summarized by a vector p from the unit simplex Δ, where p_j is the probability that a vote is cast for candidate $j \in \{1, 2, 3\}$.[4] Any idiosyncratic uncertainty is averaged out in a large electorate: if n is large then vote shares are almost always close to p.[5] Here, however, the voter is uncertain about the electoral situation: she does not know p. This (aggregate) uncertainty is captured by a continuous density $f(p)$ with full support on Δ. Others' actions are seen as only conditionally independent; unconditionally they are correlated. Taking expectations over p,

$$\Pr[x] = \int_{\triangle} \Pr[x \,|\, p] f(p)\, dp \quad \textbf{where} \quad \Pr[x \,|\, p] = \frac{n!}{x_1! x_2! x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}. \tag{1}$$

Payoffs. An instrumental voter cares about the identity of the winner.[6] Her payoff (a von Neumann-Morgenstern utility) for a win by candidate j is u_j, where $u_1 > u_2 > u_3$.

Pivotal Probabilities and Optimal Voting in Large Electorates

Here we examine the probabilities of pivotal events (in which a vote makes a difference) as the electorate size grows. We then find the instrumental voter's optimal decision rule.

Aggregate Uncertainty in Large Electorates. A voter faces two sources of uncertainty. Firstly, she is uncertain about p; this is aggregate uncertainty. Secondly, conditional on p she is uncertain of the realized decisions of others. A large electorate eliminates the second (idiosyncratic) source of uncertainty, and so beliefs are entirely determined by aggregate uncertainty.

If $n = x_1 + x_2 + x_3$ grows large then $\Pr[x \,|\, p] \propto p_1^{x_1} p_2^{x_2} p_3^{x_3}$ becomes sharply peaked around its maximum at $p = x/n$. This means that the density $f(p)$ only matters local to this maximum, and so (for large n) we can replace $f(p)$ with $f(x/n)$ in the integral of equation (1). So,

$$\Pr[x] \approx f\left(\frac{x}{n}\right) \frac{n!}{(n+2)!} \int_{\triangle} \frac{(n+2)!}{x_1! x_2! x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}\, dp = \frac{f(x/n)}{(n+2)(n+1)} \approx \frac{f(x/n)}{n^2}. \tag{2}$$

The central equality is obtained by spotting that the integrand is the density of a Dirchlet distribution, and our use of the symbol "$\approx$" is made precise in the following formal statement.

Lemma 1 (Probabilities in Large Electorates). $\lim_{n\to\infty} \max_{x \in X} |n^2 \Pr[x] - f(x/n)| = 0.$

[4]The relevant unit simplex is $\Delta = \{p \in \mathcal{R}_+^3 \,:\, p_1 + p_2 + p_3 = 1\}$. Similarly, $X = \{x \in \mathcal{Z}_+^3 \,:\, x_1 + x_2 + x_3 = n\}$.

[5]We note that p does not necessarily represent voters' true preferences, but instead captures voting intentions; the votes of others might incorporate their own strategic voting decision.

[6]Voters care only about who wins their own districts. The model of Hughes (2016) allows payoffs to depend on which government forms. His voters must consider the likely outcomes in other constituencies when deciding how to vote.

Event	$u(\text{Vote } 1)$	$u(\text{Vote } 2)$	$u(\text{Vote } 1) - u(\text{Vote } 2)$	Description
$x_1 = x_2 > x_3 + 1$	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 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 Two 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. Relevant Pivotal Events and Voter Payoffs

In large electorates, the fraction of votes cast for each candidate is close to x/n only when $p = x/n$, and so the relative probability of observing such an outcome is determined by the density $f(x/n)$. Moreover, the probability of each electoral event declines at the same rate $1/n^2$; in an IID specification, outcome probabilities go to zero at different rates as the electorate size increases.

Optimal Voting. Given that $u_1 > u_2 > u_3$, a voter optimally chooses between candidates 1 and 2. Her vote only matters when it breaks a two-way tie to determine the winner, or when it creates a tie (and so prevents an outright win) in a near-tie situation when either 1 or 2 are one vote behind the leading candidate. (Three-way ties are much less likely, and play no important role.) Focusing on these pivotal events of interest, Table 1 reports exhaustively the expected payoff from votes for the alternatives under each relevant pivotal event. The asymptotic behavior (as the electorate grows) of pivotal probabilities is characterized in our second lemma.

Lemma 2 (Tie Probabilities)**.** *The probabilities of two-way and near two-way ties satisfy*

$$\lim_{n\to\infty} n \Pr[x_i = x_j > x_k + 1] = \lim_{n\to\infty} n \Pr[x_i + 1 = x_j > x_k + 1] = p_{ij}, \tag{3}$$

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

As the proof (in the appendix) shows, equation (3) follows from the application of Lemma 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 $p_i = p_j > p_k$.

An instrumental voter optimally backs candidate 1 whenever the expected difference (between the two options) is positive; she backs candidate 2 otherwise. Allowing the electorate to grow, we can replace the probabilities of the pivotal events with the expressions p_{12}, p_{23}, and p_{13}. Doing so, we obtain the following proposition which characterizes optimal strategic voting for large n.

Proposition 1 (Optimal Voting Decision)**.** *If the electorate size is sufficiently large, then the instrumental voter should vote strategically (for her second-preference candidate) if and only if*

$$\frac{u_1 - u_3}{u_2 - u_3} < \frac{p_{23} + 2p_{12}}{p_{13} + 2p_{12}} \;\; \textit{or equivalently} \;\; \tilde{u} < \Lambda \;\; \textit{where} \;\; \tilde{u} \equiv \frac{u_1 - u_2}{u_1 - u_3} \;\; \textit{and} \;\; \Lambda \equiv \frac{p_{23} - p_{13}}{p_{23} + 2p_{12}}. \tag{4}$$

Strategic Voting Incentives. $\tilde{u} \in [0, 1]$ is a voter's allegiance to her first preference: the numerator is the benefit of her first relative to her second preference; the denominator is her desire to defeat the third candidate. Λ is positive if and only if the second choice candidate is more likely than the favorite to tie with the disliked candidate. If so then $\Lambda \in [0, 1]$ proportionally measures the incentive to vote strategically. This translates into the probability of a strategic vote: if $\tilde{u} \sim H(\cdot)$ then $\Pr[\tilde{u} < \Lambda] = H(\Lambda)$. If $\tilde{u}$ is uniformly distributed then this probability is simply Λ.

The Incentive to Vote Strategically

We now specify parametrically the form of beliefs about the underlying popularity of the candidates, and we identify the conditions under which a voter should act strategically.

Dirichlet Beliefs. The Dirichlet distribution for $f(p)$ arises naturally when a voter samples the voting intentions of others.[7] Specifically, suppose that a voter with a uniform prior over $\triangle$ observes the voting intentions of a random sample of s others. If y is a vector where y_j is the number in the sample who support j, so that $y_1 + y_2 + y_3 = s$, then $f(p \,|\, y) \propto \Pr[y \,|\, p] \propto \prod_{j=1}^{3} p_j^{y_j}$. This is the kernel of a Dirichlet distribution. For convenience, define $\pi_j = y_j / s$, the sample fraction who support j, and assemble into the vector $\pi \in \triangle$. With this formulation,

$$f(p \,|\, y) \propto \tilde{f}(p) = \left[\prod\nolimits_{j=1}^{3} p_j^{\pi_j}\right]^s. \tag{5}$$

$\pi = \arg\max \tilde{f}(p)$ is the modal belief, while s indexes the precision of beliefs around this mode.

In the Dirichlet case, the asymptotic probabilities of two-way ties and near two-way ties in large electorates (these are the terms $p_{ij} = \lim_{n\to\infty} n \Pr[x_i = x_j > x_k]$ from Lemma 2) satisfy

$$p_{ij} \propto G(\pi_k, s) \quad \text{where} \quad G(\varpi, s) \equiv \int_{1/3}^{1/2} [z^{1-\varpi}(1 - 2z)^{\varpi}]^s \, dz, \tag{6}$$

where i, j, and k are three different candidates. Equation (6) gives us the next result.

[7] McKelvey and Ordeshook (1972) and Hoffman (1982) defined beliefs directly over X, using a continuous approximating density. The analogue here is the density $f(p)$. They specified uniform and Gaussian kernels.

Lemma 3 (Strategic Incentives with Dirichlet Beliefs)**.** *For Dirichlet beliefs,*

$$\Lambda = \frac{G(\pi_1, s) - G(\pi_2, s)}{G(\pi_1, s) + 2G(\pi_3, s)}, \tag{7}$$

where $G(\varpi, s) \equiv \int_{1/3}^{1/2} [z^{1-\varpi}(1-2z)^{\varpi}]^s \, dz$. $G(\varpi, s)$ *is a decreasing function of* ϖ*, and satisfies* $G(\varpi, s) = 2^{\omega s}\,\mathrm{Beta}(\frac{1}{3}; \varpi s + 1, (1-\varpi)s + 1)/2^{s+1}$ *where* $\mathrm{Beta}(t; a, b)$ *is the incomplete beta function.*

The strategic incentive Λ forms the basis for our comparative-static predictions.[8] Before deriving those predictions, however, we find the situations in which that incentive is positive.

Risk Population. A voter should consider a strategic vote if and only if her favorite is less likely than her second preference to tie with the disliked candidate. In a large electorate this is true (from Lemma 3) if and only if $G(\pi_2, s) < G(\pi_1, s)$, which holds if and only if $\pi_2 > \pi_1$. Note that this does not require a voter's favorite to be in (expected) third place.

Proposition 2 (Risk Population)**.** *Under the Dirichlet specification, and in a large electorate, a voter faces a positive incentive to vote strategically if and only if she expects her favorite candidate to trail behind her second-preference candidate. That is,* $\Lambda > 0$ *if and only if* $\pi_1 < \pi_2$*.*

Fisher (2004) and Kselman and Niou (2010) also claimed that a strategic vote only makes sense if a second choice is expected to place ahead of a first choice, but did not offer a proof.[9]

The Precision of Beliefs. The incentive to vote strategically is present even if a voter's favorite is in a strong position, so that $\pi_2 > \pi_1 > \pi_3$. Nevertheless, the incentive is weaker in this case, and stronger when $\pi_1 < \{\pi_2, \pi_3\}$. This is very sharply true when the precision of beliefs is high: a tie involving the two most popular candidates becomes much more likely than any other tie, and so the incentive becomes either very strong or very weak.

Proposition 3 (The Effect of Precise Beliefs)**.** *If* $\pi_2 > \pi_1 > \pi_3$*, so that the voter's favorite is expected to place amongst the leading pair, then* $\lim_{s\to\infty} \Lambda = 0$*, and so if* s *is large then she always votes sincerely. However, if* $\pi_1 < \min\{\pi_2, \pi_3\}$*, so that her favorite is expected to place last, then* $\lim_{s\to\infty} \Lambda = 1$*, and so if* s *is large then she always votes strategically.*

---

[8] A transformation of Λ was also reported by Fisher and Myatt (2002a) and used empirically by Herrmann (2010).

[9] Fisher (2004) referenced unpublished research which underpinned the work (Fisher and Myatt, 2002a) upon which this paper builds. Kselman and Niou (2010) also observed, correctly, that others (Ordeshook and Zeng, 1997; Alvarez, Boehmke, and Nagler, 2006) restricted to situations in a which first choice is expected to trail last, but that there is no proper basis for this. In general some structure on beliefs (such as the Dirichlet specification used here) is needed for claims about the nature of the risk population to be made.

Comparative-Static Predictions

We now turn attention to comparative-static predictions. We mainly focus on the traditional scenario in which a voter's favorite candidate is expected to run third in the district.

Shifting Popularity. We cannot isolate a single candidate's popularity: any loss in support for a candidate is a gain for others. Nevertheless, we note, via equation (7), that the strategic incentive is decreasing in the popularity of the voter's favorite, but increasing in the popularities of the others. This pins down the effect of moving support to or from the voter's favorite.

Proposition 4 (The Effect of Shifting Candidate Popularity)**.** *A shift in popularity from the voter's favorite candidate toward one of the others increases the incentive to vote strategically.*

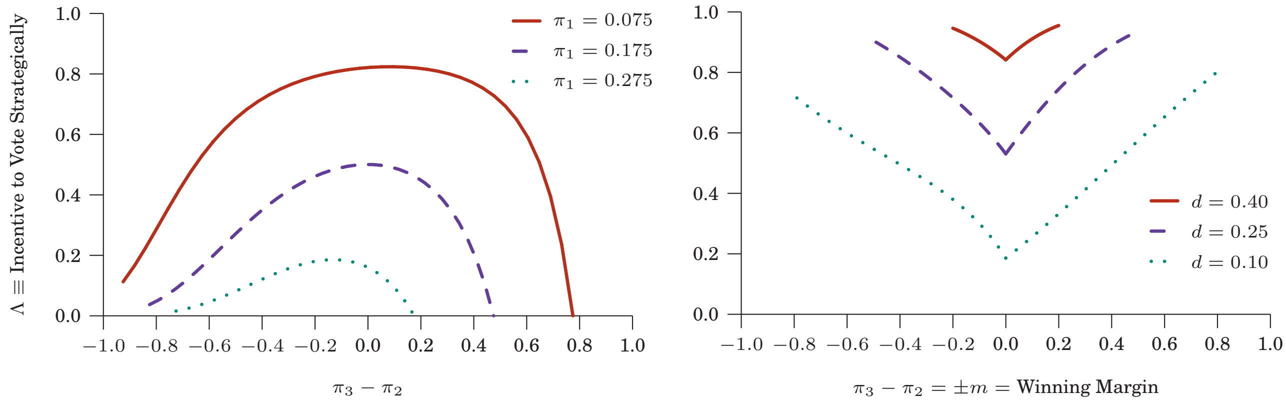
Movements between the second and third candidates are more involved. Fixing π_1, the response to $\pi_2 - \pi_3$ is non-monotonic. The incentive is greater when π_2 and π_3 are not too far apart (Figure 1a), but it is not maximized at $\pi_2 - \pi_3 = 0$. This conflicts with the common intuition that strategic voting should be greatest in a close race between the second and third choices.

Distance from Contention and the Winning Margin. According to many, strategic voting should be attractive to voters who are indifferent between their first and second preferences, but strongly dislike others (Heath, Curtice, Jowell, Evans, Field, and Witherspoon, 1991). It is also thought that a voter will "avoid wasting her vote" if the difference in support between the second-placed candidate and her third-placed favorite increases. Finally, many have suggested that strategic voting should be greater in close races because ties are more likely (Cain, 1978; Niemi, Whitten, and Franklin, 1992; Evans, 1994; Fieldhouse, Pattie, and Johnston, 1996; Cox, 1997). Recent empirical evidence (Kawai and Watanabe, 2013) also favors this final suggestion.[10]

The importance of the two key district characteristics (distance and margin) suggests that a careful theoretical study of them is warranted. Here, we isolate the effects of these two parameters.

Firstly, we consider a situation in which $\pi_3 > \pi_2 > \pi_1$, so that the disliked third candidate is expected to lead. In this context, we can define the margin of victory of the leading candidate and the distance from contention of the voter's favorite as $m = \pi_3 - \pi_2$ and $d = \pi_2 - \pi_1$. Any shift in support between two candidates influences both of these measures.

[10]The distance from contention has been used to test strategic-voting measures (Niemi, Whitten, and Franklin, 1992; Evans and Heath, 1993; Niemi, Whitten, and Franklin, 1993; Franklin, Niemi, and Whitten, 1994; Heath and Evans, 1994). Curtice and Steed (1988, 1992, 1997) used the victory margin to infer strategic voting from aggregate data.



(A) Shifting Between Candidates 2 and 3 (B) The Effect of the Winning Margin

Notes. The precision of voters' Dirichlet beliefs is set at $s = 10$. Recall that the strategic incentive Λ coincides with the probability of a strategic vote when $\tilde{u} \sim U[0,1]$. Fixing π_1, Figure 1a relates $\pi_3 - \pi_2$ to Λ. The right-hand extreme of each line corresponds to $\pi_2 = \pi_1$; the left-hand extreme corresponds to $\pi_3 = 0$. Fixing the distance from contention $d = \min\{\pi_2, \pi_3\} - \pi_1$, Figure 1b relates the winning margin to the strategic incentive.

FIGURE 1. The Effect of the Electoral Situation on Strategic Incentives

To make progress, we solve for π_1, π_2, and π_3 in terms of m and d via the constraint $\pi_1+\pi_2+\pi_3 = 1$:

$$\pi_1 = \frac{1-m-2d}{3}, \quad \pi_2 = \frac{1+d-m}{3}, \quad \text{and } \pi_3 = \frac{1+2m+d}{3}. \tag{8}$$

We also consider situations in which $\pi_2 > \pi_3 > \pi_1$, so that the voter's second-preference candidate has greater expected popularity. The definitions of m and d become $m = \pi_2 - \pi_3$ and $d = \pi_3 - \pi_1$, and the solutions for the candidates' popularities become

$$\pi_1 = \frac{1-m-2d}{3}, \quad \pi_2 = \frac{1+2m+d}{3}, \quad \text{and } \pi_3 = \frac{1+d-m}{3}. \tag{9}$$

Inspecting equations (8) and (9), and fixing the margin of victory, an increase in the distance from contention lowers the popularity of the voter's favorite, and pushes support equally toward the other two. From Proposition 4, the incentive to vote strategically must rise.

A wider margin of victory (for a fixed distance from contention) is more complex. It pushes the favorite down, which increases the strategic-voting incentive; however, the other candidates move in opposite directions, which has an ambiguous effect. Proposition 5 finds the overall effect.

Proposition 5 (Margin of Victory and Distance from Contention)**.** *Fixing the expected margin of victory, the strategic-voting incentive is increasing in the distance from contention. Fixing the distance from contention, the incentive is increasing in the expected margin of victory.*

A common informal argument is that voters should act strategically when their probability of influence is high. However, what should matter is the relative probability of pivotal events. Fixing the distance from contention, an increased margin of victory might lower the probability of a tie for the lead, but the probability that a tie (if it occurs) involves the leading two candidates rises. So, the incentive to vote strategically increases with the margin of victory after controlling for the distance from contention. If strategic voting happens in close races then this may be because the distance from contention of the favorite is large, rather than the narrow margin.

A further observation is that the effect of the margin of victory seems (from Figure 1b) to be weaker than that of the distance from contention. We confirm this formally here.

Proposition 6 (Relative Comparative-Static Effects)**.** *The distance from contention has a stronger effect than the margin of victory: $\partial\Lambda/\partial d > \partial\Lambda/\partial m$. Moreover, if $\pi_3 > \pi_2 > \pi_1$ then $\partial\Lambda/\partial m \to 0$ as $d \to 0$: if the disliked candidate is ahead and if the distance from contention is small then changes in the expected margin of victory have a negligible effect.*

Strategic Voting in Equilibrium: Classic Theories

Our results so far do not consider explicitly the process via which others' votes arise.[11] Here (in this section) we review the insights offered by some classic game-theoretic models before (in the following section) we study a game-theoretic model that incorporates aggregate uncertainty.

Duvergerian Equilibria. Many game-theoretic models involve IID voter types and decisions. In this context, suppose that play is such that $p_i \geq p_j > p_k$. In a large electorate, any tie (if a tie occurs) is almost always between i and j and so a voter should abandon $k \notin \{i, j\}$. Very little game-theoretic reasoning is needed. Beginning from a situation in which one candidate is believed to hold third place, a single step of best replies takes voters to a Duvergerian world.

In the presence of aggregate uncertainty, other kinds of ties remain possible. Nevertheless, the argument holds in spirit if beliefs are precise: from Proposition 3, the incentive to abandon a third-placed (in expectation) candidate becomes very strong as a voter's pessimism solidifies.

If a voter is less certain (if s is not so large) then the third candidate can retain some support. However, game-theoretic reasoning now kicks in: that loss of support (a reduction in π_1, via

[11] Nevertheless, Lemmas 1–2 and Proposition 1 do apply in the context of a game-theoretic analysis. $f(p)$ represents a voter's beliefs about voting probabilities, and not the probabilities of sincere preference orderings. Given that these beliefs take into account the strategies of others, they continue to underpin a best-responding voter's behavior.

Proposition 4, increases the incentive to vote strategically) further increases strategic voting, and a few steps of best replies lead to a Duvergerian equilibrium. If the positions of the candidates are commonly understood then it is difficult to obtain an equilibrium with three-candidate support. Moreover, a Duvergerian equilibrium only requires voters to achieve a common understanding of which two candidates are the contenders; exact knowledge of p is not needed.

Non-Duvergerian Equilibria. The pressure toward a Duvergerian outcome requires the third candidate to fall behind. However, if a tie for second place is constructed then the "non-Duvergerian" equilibria that were emphasized by Cox (1994, 1997) may be considered.

Within an IID framework, suppose that $\bar{p}_3 > \bar{p}_2 > \bar{p}_1$ where $\bar{p}_j$ is the probability that another voter truly prefers candidate j.[12] Suppose that supporters of candidate 3 vote sincerely, so that $p_3 = \bar{p}_3$, and that others dislike candidate 3. If a (delicately calculated) fraction of those who prefer candidate 2 switch to (the less popular!) candidate 1 so that $p_1 \approx p_2$ then both candidates 1 and 2 may contend for the lead. This construction requires the supporters of a popular candidate to have exact common knowledge of $\bar{p}$ so that a precise fraction can switch in the wrong direction. The equilibrium is also unstable: any small deviation results in a large strategic incentive for supporters of the trailing candidate, so leading back to a Duvergerian outcome (Fey, 1997).

The intuition offered by Cox (1997, p. 86) is likely correct, but differs technically from what is required for a non-Duvergerian equilibrium. He considered the Ross and Cromarty constituency in the 1970 British General Election. A split between the Liberal and Labour candidates permitted a Conservative win. His interpretation was that "it was not clear who was in third and who in second." However, non-Duvergerian equilibria require precise common knowledge of the true popularities of all candidates; there is no room for any lack of clarity.

Classic Predictions. Summarizing, Cox (1997) has suggested that districts should be clustered in two groups: those in which the top two candidates receive nearly all the votes, and those in which the second and third placed candidates are close. Cox (1997, pp. 86–89) evaluated this prediction using the SF ratio.[13] This is the ratio of the vote share for the second loser (S) to that of the first loser (F). Taken literally, the Duvergerian and non-Duvergerian equilibria of IID models would generate SF ratios of zero and one. This is the bi-modality hypothesis.

[12] Here we are no longer labeling candidates according to the preference order of a single voter.

[13] Although not used extensively, the SF ratio and the associated bi-modality hypothesis do continue to appear more recently in the literature. See, for example, more recent work by Moser and Scheiner (2009) and Selb (2012).

Strategic Voting in Equilibrium: A Newer Theory

More recent equilibrium theories of voting have allowed for aggregate uncertainty. Here we extend the model of Myatt (2007) to allow for the (empirically important) possibility of sincere voters, and we show that our predictions (Propositions 5 and 6) hold in a game-theoretic world.

A Model of Strategic Voter Coordination. We simplify by fixing the proportion of voters who prefer and vote for candidate 3 at π_3, where $\frac{1}{3} < \pi_3 < \frac{1}{2}$. Candidate 3 never trails in third place, but loses if there is coordination behind a challenger. Other voters' preferences satisfy $\min\{u_1, u_2\} > u_3$ and, without loss of generality, we set $u_3 = 0$. These other voters wish to stop candidate 3, but need to coordinate. They vary in their relative preferences, and also in their information about others. There is no common knowledge of the best-placed challenger.

The district situation is summarized by a variable θ which determines, via equation (10), the popularity of candidate 1 versus candidate 2. Voters begin with an (improper) flat prior over θ.

A voter's type combines: her preferences, determined by $\log(u_1/u_2)$; an informative signal $\hat{\theta}$ of θ; and her willingness to act instrumentally. Two key parameters are the standard deviation $\sigma > 0$ of voters' preference intensities, and the precision $\alpha \geq 1$ of a voter's signal. Specifically,

$$\hat{\theta} = \theta + \eta \text{ and } \frac{\log(u_1/u_2)}{\sigma} = \theta + \eta + \varepsilon, \quad \text{where} \quad \eta \sim N\left(0, \frac{1}{\alpha}\right) \text{ and } \varepsilon \sim N\left(0, \frac{\alpha - 1}{\alpha}\right), \qquad (10)$$

where η and ε are uncorrelated with each other and across the electorate. A voter is instrumentally motivated with probability ψ; otherwise she votes exogenously for her first preference.

Firstly, notice that $\sigma^{-1}\log(u_1/u_2) \sim N(\theta, 1)$ and so $\Pr[u_1 > u_2] = \Phi(\theta)$ where $\Phi(\cdot)$ is the distribution function of the standard normal. Hence θ (which is unknown to voters) determines the fraction of voters (amongst those who dislike candidate 3) who prefer candidate 1 to candidate 2.

Secondly, a voter's updated beliefs satisfy $\theta \mid \hat{\theta} \sim N(\hat{\theta}, 1/\alpha)$. Beliefs differ, owing to variation in $\hat{\theta}$; however, if the precision α grows then voters begin to agree on the best challenger to candidate 3. We have imposed a lower bound: $\alpha \geq 1$. This is because a voter can always use her own preference as an informative signal of θ.[14] (The case $\alpha = 1$ implies that $\log(u_1/u_2) = \hat{\theta}$.)

Finally, $\sigma^2 = \operatorname{var}[\log(u_1/u_2) \mid \theta]$ is the variability in the intensity of voters' preferences: if σ^2 is larger then more voters care relatively more deeply about their favorite candidate.

[14] A voter's preferences provide no information about θ beyond the signal $\hat{\theta}$. This is without loss of generality; if this were not the case, then we could redefine $\hat{\theta}$ appropriately, as explained by Myatt (2007, Section 2.1.3, p. 259).

The Negative Feedback Effect. Given her type, a rise in $\log(u_1/u_2)$ pushes a voter to support candidate 1. However, a fall in $\hat{\theta}$ indicates that candidate 2 may be a better challenger, and so a strategic vote might be warranted. Hence, if $\log(u_1/u_2)$ and $\hat{\theta}$ differ in sign then a voter may consider a strategic vote. Here we consider a simple functional form for resolving the trade-off between a voter's preference and her signal. Suppose that she adopts the strategy:

$$\text{Vote 1} \quad \Leftrightarrow \quad \log\left(\frac{u_1}{u_2}\right) + b\hat{\theta} \geq 0, \tag{11}$$

where $b > 0$ is a voter's willingness to respond strategically to her information about others' preferences. The properties of the normal distribution mean that a linear form is natural. This is because in this setting, a voter in a large electorate should vote for candidate 1 if and only if

$$\log\left(\frac{u_1}{u_2}\right) + \lambda \geq 0 \quad \text{where} \quad \lambda = \lim_{n\to\infty} \log \frac{\Pr[x_1 = x_3 > x_2 \,|\, \hat{\theta}]}{\Pr[x_2 = x_3 > x_1 \,|\, \hat{\theta}]}. \tag{12}$$

In this expression λ is the log likelihood ratio of pivotal events. The normal specification leads to log likelihood ratios that are linear in the mean of a voter's beliefs about the variable of interest.[15]

Lemma 4 (Optimal Reaction to Information about the Electoral Situation). *If other instrumental voters (amongst those who dislike candidate 3) adopt a linear voting strategy by voting for candidate 1 if and only if* $\log(u_1/u_2) + b\hat{\theta} \geq 0$*, then for an instrumental voter with signal* $\hat{\theta}$*,*

$$\lim_{n\to\infty} \log \frac{\Pr[x_1 = x_3 > x_2 \,|\, \hat{\theta}]}{\Pr[x_2 = x_3 > x_1 \,|\, \hat{\theta}]} = \hat{b}\hat{\theta} \quad \textit{where} \quad \hat{b} = B(b; \alpha, \sigma^2, \pi_3, \psi), \tag{13}$$

for some function $B(b; \alpha, \sigma^2, \pi_3, \psi)$*. This function is increasing in the coordination* π_3 *required to defeat the disliked candidate; in the precision* α *of a voter's information; and in the heterogeneity* σ^2 *of preferences. It is decreasing in the response* b *of others to their signals, and decreasing in the proportion* ψ *of instrumental voters. If* $\psi = 1$*, so that everyone is instrumental, then*

$$\hat{b} = B(b; \alpha, \sigma^2, \pi_3, 1) = 2\Phi^{-1}\left(\frac{\pi_3}{1-\pi_3}\right)\sqrt{\alpha + \frac{\alpha(\alpha-1)\sigma^2}{(\sigma+b)^2}}. \tag{14}$$

If $\log(u_1/u_2) + \hat{b}\hat{\theta} > 0$ *and if the electorate size is sufficiently large then a voter optimally votes for candidate 1; if the opposite strict inequality holds then she should vote for candidate 2.*

[15] More generally: if others react monotonically to their signals, then a voter's best reply takes the form: vote for candidate 1 if $\log(u_1/u_2) + a + b\hat{\theta} \geq 0$. When the prior is neutral or imprecise, then the unique stable equilibrium entails $a = 0$. All of the claims in this paper are robust to the inclusion of a non-neutral informative prior or the inclusion of public signals about the candidates' popularities, so long as private signals are sufficiently precise.

If others place greater emphasis on their signals then a small underlying asymmetry in favor of a candidate enables that candidate to tie, simply because that asymmetry is amplified by strategic voting. This means that the states of the world that a voter compares when working out the odds of pivotal events become close, and this weakens the strategic-voting incentive.

The negative feedback effect also explains why a voter's reaction to her signal is increasing in the heterogeneity of others' preferences. An increase in σ^2 makes others less willing to abandon their first choices. This pushes apart the values of θ that induce tied outcomes, and so (for a fixed signal realization) it increases the relative likelihood of one pivotal event compared to the other.

Equilibrium. We define a voting equilibrium as a positive coefficient $b^\star$ such that if all other instrumental voters use it (they vote for candidate 1 if and only if $\log(u_1/u_2) + b^\star\hat{\theta} > 0$) then a voter finds it optimal to do the same if the electorate is sufficiently large.[16] To find such an equilibrium we seek a coefficient $b^\star$ satisfying $b^\star = B(b^\star; \alpha, \sigma^2, \pi_3, \psi)$. We know (from Lemma 4) that $B(\cdot)$ is decreasing in b, and so $b^\star = B(b^\star; \alpha, \sigma^2, \pi_3, \psi)$ has a unique positive solution.

Proposition 7 (Equilibrium with Privately Informed Voters)**.** *There is a unique voting equilibrium. The coefficient $b^\star$ is increasing in the coordination π_3 required to defeat the disliked candidate, in the precision α of information, and in the heterogeneity σ^2 of preferences. It is decreasing in the proportion ψ of instrumental voters. $\lim_{\alpha\to\infty} b^\star = \infty$, and so very well informed voters support whomever they see as the more popular challenger.*

Voters do not fully coordinate. A Duvergerian outcome requires voters to agree which challenger is stronger, and this common understanding is only achieved as $\alpha \to \infty$. Away from the limit (when information is imperfect) the outcome is only partially Duvergerian: in some cases (but not always) a voter switches to someone whom she perceives as more popular.[17] Interestingly, the introduction of sincere voters (that is, $\psi < 1$) increases the willingness of instrumental voters to react to their signals of the electoral situation. This dampens the reduction in strategic voting that might otherwise be expected from an increase in sincere voting; this is a new finding.

---

[16] This solution concept was defined by Myatt (2007) and has been applied with models of party leadership (Dewan and Myatt, 2007), costly voter turnout (Myatt, 2015) and protest voting (Myatt, 2016). It is defined for a sequence of voting games indexed by an electorate size n. A conventional approach would be to find an equilibrium for each n, and to take the limit as $n \to \infty$. Such a limit corresponds to the equilibrium described here.

[17] This is very different from the non-Duvergerian construction (Cox, 1994, 1997) in which some voters switch to a less popular candidate; there is no relationship between that and the equilibrium of Proposition 7.

Comparative-Static Predictions. A voter is more willing to vote strategically when her information is good and when others' preferences are more extreme (Proposition 7). Here we consider the electoral situation. Previously, we summarized this via the Dirichlet parameters π_1, π_2, and π_3. Here, those parameters refer to the true underlying popularity of the candidates. Hence

$$\frac{\pi_1}{1-\pi_3} = \Pr[u_1 > u_2] = \Phi(\theta) \quad \Leftrightarrow \quad \theta = \Phi^{-1}\left(\frac{\pi_1}{1-\pi_3}\right). \tag{15}$$

We focus our exposition on a situation in which $\pi_1 < \pi_2 < \pi_3$, which implies that $\theta < 0$.

The equilibrium coefficient $b^\star$ depends on the electoral situation only via π_3. From equation (8), π_3 is increasing in both the distance from contention and margin of victory parameters. Straightforwardly, then, here our comparative-static predictions are maintained.

The typical incentive to vote strategically depends not only on the equilibrium coefficient $b^\star$ but also upon the average signal received by a voter. On average, that signal is equal to θ. Thus, the average incentive (in terms of λ) to vote strategically for candidate 2 is

$$\mathrm{E}[\lambda \mid \theta] = -b^\star\theta = -b^\star\Phi^{-1}\left(\frac{\pi_1}{1-\pi_3}\right) = b^\star\Phi^{-1}\left(\frac{\pi_2}{1-\pi_3}\right), \tag{16}$$

where the final equality is obtained from the symmetry properties of $\Phi(\cdot)$. It is straightforward to show that $\pi_2/(1-\pi_3)$ is increasing in the distance from contention d and the margin of victory m. Hence, the comparative-static predictions of Proposition 5 apply here too. Moreover, we can also evaluate the relative size of the margin and distance effects.[18]

Proposition 8 (Margin of Victory and Distance from Contention in a Voting Equilibrium)**.** *The results of Propositions 5–6 apply: fixing the expected margin of victory, the strategic-voting incentive is increasing in the distance from contention; and, fixing the distance from contention, the incentive is increasing in the expected margin of victory. Moreover,* $\partial\,\mathrm{E}[\lambda \mid \theta]/\partial m \to 0$ *as* $d \to 0$*; the marginality effect is negligible when the distance from contention is small.*

In summary, our game-theoretic analysis supports our (decision-theoretic) predictions: the marginality effect should be weak and opposite to that expected by Cain (1978) and others, and we do not expect the bi-modality of the SF ratio suggested by Cox (1997).

[18]Proposition 8 extends results from Myatt (2007) by noting that the marginality effect is substantially weaker than the distance effect (this partially replicates Proposition 6) in districts where the challengers are close.

UK General Elections. The United Kingdom operates the plurality rule throughout its parliamentary constituencies. Northern Ireland is unique; elections are dominated by region-specific parties. Scotland and Wales have successful nationalist parties, and so four-party competition.

We focus, therefore, on England, which accounts for 82% of the UK's 650 constituencies. Throughout six recent general elections (1987, 1992, 1997, 2001, 2005, and 2010) the three major parties competed in almost every constituency.[19] In all but a very small handful of cases, they occupied the first, second, and third positions at the constituency level. For the most recent general election, held in 2015, the rise of the UK Independence Party resulted in genuine four-party competition throughout England, and so this election is not considered here.

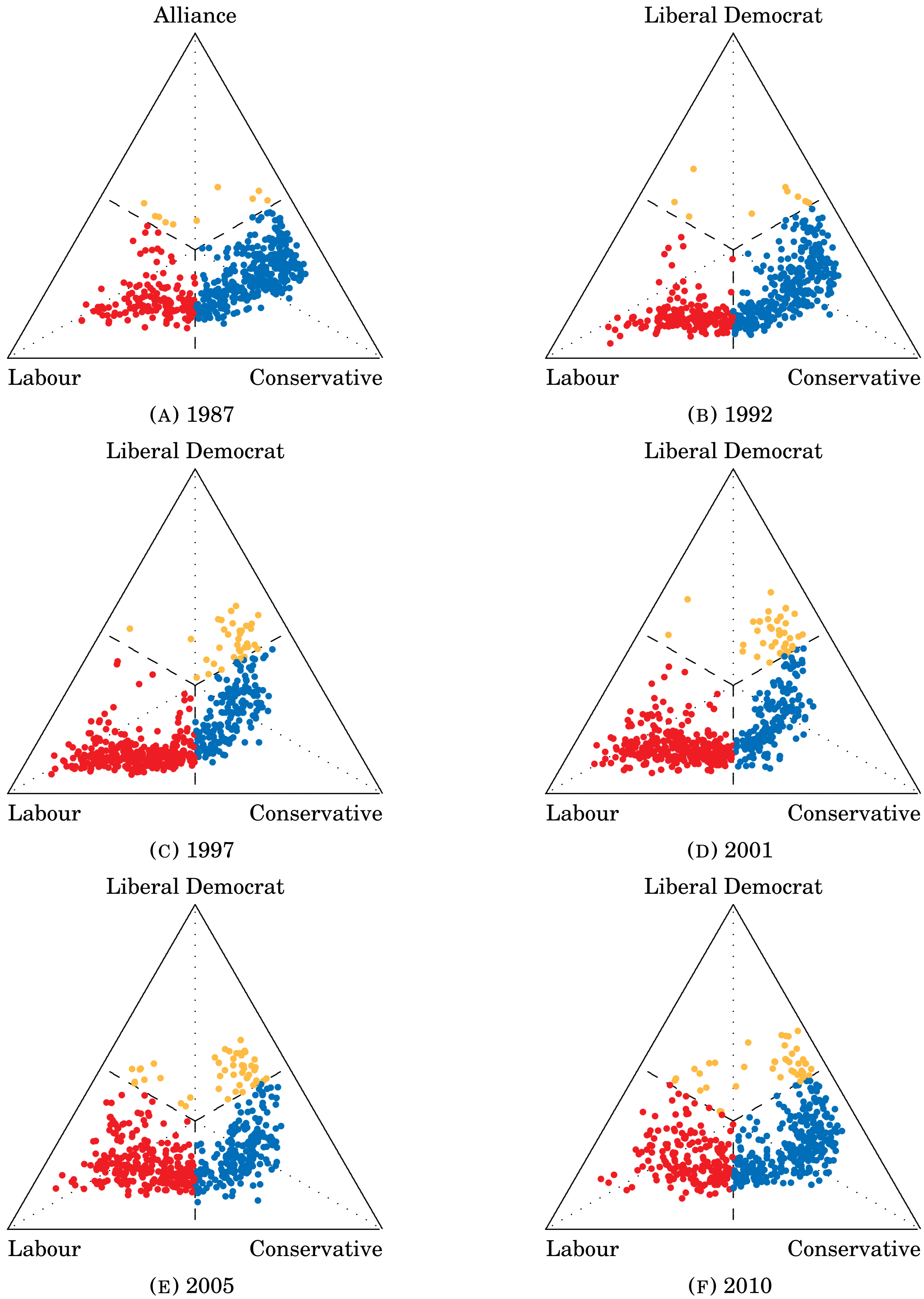
Given the three-party system throughout England in this period, it is convenient to represent the vote shares via the simplex plots of Figure 2. The corners of each triangle represent a 100% vote share (amongst the major parties) for one party. This triangle is separated into three separate win zones by the hatched lines; those hatched lines represent ties for the lead, whereas the dotted lines mark ties for second place. A parliamentary constituency is a bullet ("•") at a weighted average of the three corners of the simplex, where the weights are the vote shares.

Classic Equilibrium Predictions. For classic theories of strategic voting, a Duvergerian equilibrium (only two candidates receive support) is a clear prediction. Moreover, outcomes that are close to Duvergerian are also consistent with variations of such theories. For example, if some voters are exogenously sincere then instrumental voters continue to vote only for the two leading candidates, whereas there will be some third-candidate support from sincere voters.

We have noted that the theoretical support for a non-Duvergerian equilibrium is less clear: supporters of a more popular candidate must move to a trailing candidate and in very precise numbers. An exact tie for second place is needed even in the presence of exogenously sincere voters; any move away from an exact second-place tie is inconsistent with this kind of equilibrium.

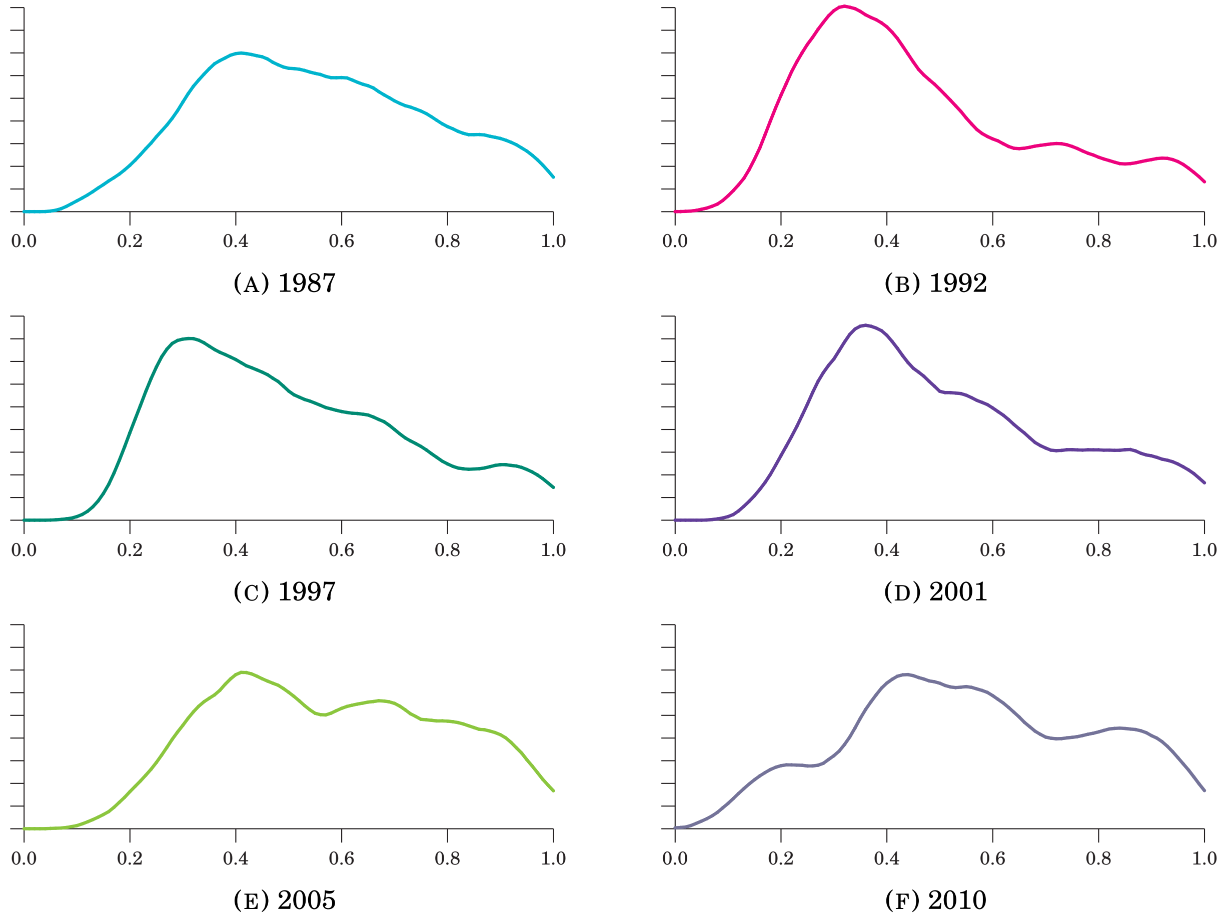
Looking at Figure 2, there are no purely Duvergerian constituencies. Nevertheless, a Duvergerian equilibrium amongst instrumental voters is consistent with points that are close to the solid boundaries of each simplex. Non-Duvergerian equilibria would cluster around the dotted lines. Again, there are none; but this time, exact clustering is necessary for such equilibria.

[19]The third party in 1987 was the SDP-Liberal Alliance; they subsequently merged to become the Liberal Democrats.



Notes. Each constituency is illustrated by a bullet point ("•") at the 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. The hatched lines separate the parties' win zones. The dotted lines correspond to ties for second place.

FIGURE 2. Simplex Plots of English General Election Results



Notes. These density plots use the Epanechnikov kernel with a bandwidth of $h = 0.1$

FIGURE 3. Kernel Density Plots of the SF Ratio for English Constituencies

The Bi-Modality Hypothesis. The theoretical and empirical support for second-place ties is questionable. Nevertheless, non-Duvergerian equilibria lie behind the bi-modality hypothesis. The SF ratio is the vote share of the second loser relative to the first loser. SF $= 0$ for an exactly Duvergerian outcome, and SF $= 1$ for a non-Duvergerian equilibrium. In an important book, Cox (1997) looked for bi-modality of the SF ratio in British constituencies. However, there are problems with sample selection, the use of marginality, and the lack of a null-hypothesis.

Firstly, Cox (1997, pp. 88–89) remarked that "the constituencies chosen for inclusion in the analysis were those in which it would have made sense for voters to consider a tactical vote" and yet he restricted attention to constituencies where Labour was third. However, there are third-party supporters (and so a strategic incentive) in all constituencies.[20] Secondly, Cox (1997) used

[20]Cox (1997) did not argue that the risk population is greater in the constituencies he has chosen; it is impossible to do so without survey data because some measure of party support distinct from the share of the vote is required.

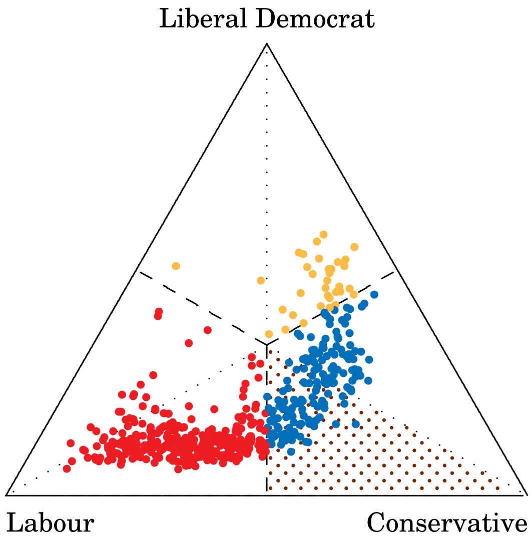
the marginality hypothesis (this plays no part in the bi-modality theory) to select restricted subsamples: he suggested that strategic voting should be relevant only when the margin of victory is slim. More precisely, he was able to plot a bi-modal distribution for the SF ratio only for marginal constituencies in which Labour came third. Thirdly, there was no null hypothesis; this is important in order to know what would constitute evidence against the bi-modality hypothesis. It requires a prediction of what constituency-level results would be if everyone voted sincerely. With the available survey data, it is impossible to pin down "sincere" at the constituency level.

While we can do very little about the absence of an effective null hypothesis, we can remove the aforementioned sample restrictions. The observed SF ratios are presented as kernel density plots in Figure 3; we see no evidence for the bi-modality hypothesis.

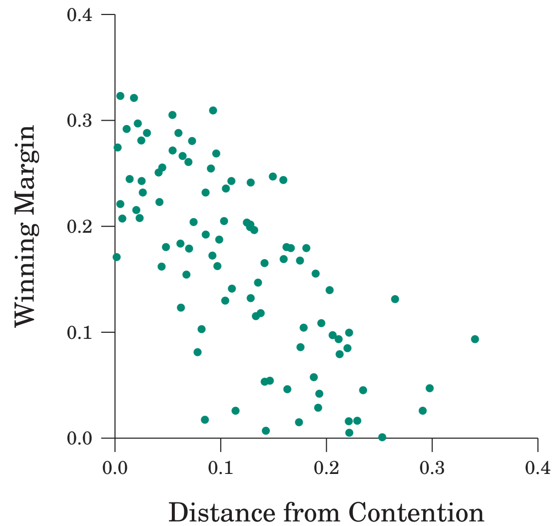
Theoretical Strategic Incentives in England. The pattern of constituency-level outcomes in England is consistent with incentives to vote strategically that are not overwhelming; otherwise we would see fully Duvergerian outcomes. So: how strong (in theory) are those strategic-voting incentives? To provide an illustrative answer we have calculated (for various parameter values of our Dirichlet model) the incentive faced by Liberal Democrat supporters (other voter types yield similar results) to switch to Labour in constituencies where the expected candidate ranking (amongst other voters) is Conservative > Labour > Liberal Democrat.

Specifically, we allocate the labels $i \in \{1, 2, 3\}$ to the Liberal Democrat, Labour, and Conservative candidates respectively. Given that $u_1 > u_2 > u_3$, this is a Liberal Democrat ($i = 1$) supporter who wishes to prevent a Conservative ($i = 3$) win. We use our Dirichlet voter-uncertainty model: beliefs about the behavior of others are described by the density $f(p) \propto (p_1^{\pi_1} p_2^{\pi_2} p_3^{\pi_3})^s$ with parameters π_1, π_2, π_3, and s. Recall that π is the modal vector of voting probabilities, and s indexes the precision of the voter's beliefs. We consider the classic situation in which a voter expects her preferred candidate to trail in third place, and dislikes the perceived leader: $\pi_1 < \pi_2 < \pi_3$.

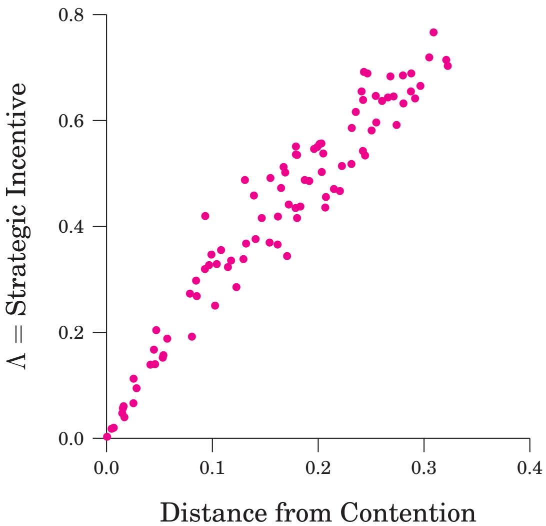
We proxy for the modal voting probabilities with realized vote shares in the 1997 General Election; similar findings emerge from other years. There were 92 constituencies of the relevant kind, and they correspond to the shaded area of Figure 4a. The strategic incentive Λ depends on the (perceived) distance from contention $d = \pi_2 - \pi_1$ and the winning margin $m = \pi_3 - \pi_2$. These variables (Figure 4b) are negatively correlated: closer races (low values of m) tend to be those where the preferred candidate is far from contention (high values of d).



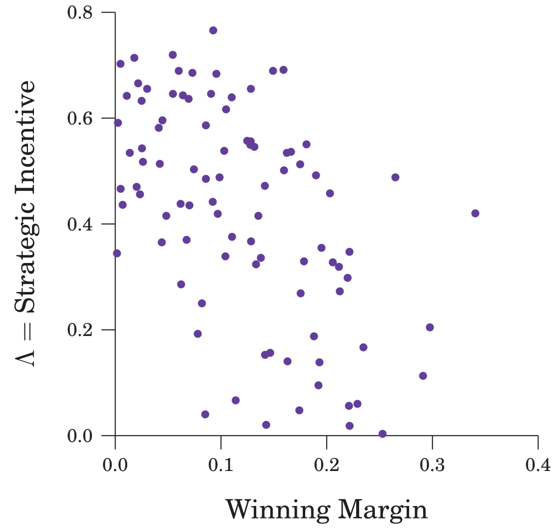
(A) Simplex Plot from Figure 2



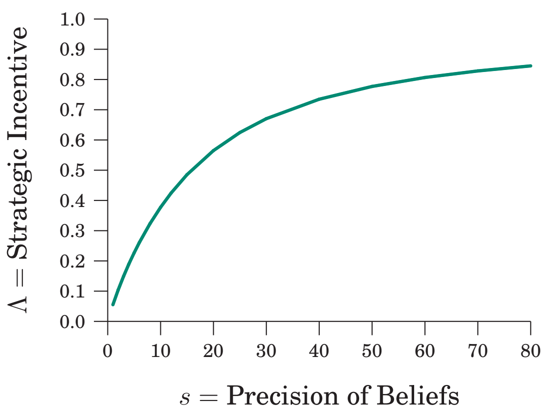
(B) Distance and Margin for Con > Lab > LD



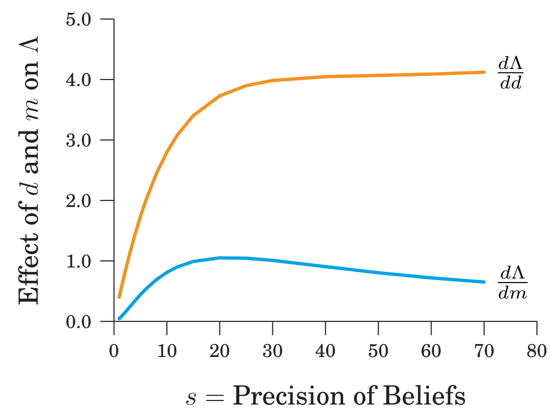
(C) Strategic Incentive vs. Distance ($s = 10$)



(D) Strategic Incentive vs. Margin ($s = 10$)



(E) Strategic Incentive vs. Belief Precision



(F) Effect of Increased Distance and Margin

Notes. English parliamentary constituencies from the General Election of 1997 in which the three major parties finished in the order Conservative > Labour > Liberal Democrat.

FIGURE 4. Liberal Democrats as Third-Placed Candidates in 1997

For these constituencies we calculated the strategic incentive for a Liberal Democrat voter to switch to Labour. From Proposition 1, she should vote strategically if and only if $\tilde{u} < \Lambda$, where

$$\tilde{u} = \frac{u_1 - u_2}{u_1 - u_3} \quad \text{and} \quad \Lambda = \frac{G(\pi_1, s) - G(\pi_2, s)}{G(\pi_1, s) + 2G(\pi_3, s)}, \tag{17}$$

where Λ is from Lemma 3 and where $G(\pi_i, s) \propto \int_{1/3}^{1/2} [z^{1-\pi_i}(1-2z)^{\pi_i}]^s \, dz$. Recall that $\tilde{u} \in [0,1]$ measures the strength of first-preference, and so Λ measures the strength of the strategic incentive. As Figure 4c (which uses $s = 10$) shows, this incentive is highly correlated with the distance from contention. Hence, taken on its own, this distance should be a strong predictor. The incentive is rather more weakly (and negatively) related to the winning margin (Figure 4d). The mild negative correlation it to be expected. Whereas the theoretical partial impact of the winning margin is positive, a higher margin is inevitably associated with a lower distance from contention. The strongly powerful effect of that distance from contention means that the lower distance mutes incentives enough to more than offset any effect of a wider winning margin.

Figures 4c and 4d are constructed for various values of the belief precision s. Figure 4e, which reports the average strategic incentive across the constituencies, verifies Proposition 3: the strategic incentive rises as a voter grows confident of the electoral situation.[21] For example, a moderately low value of $s = 10$ yields $\Lambda \approx 0.38$, whereas an intermediate value $s = 40$ generates $\Lambda \approx 0.73$. Recall that the strategic incentive is the proportion of at-risk voters who vote strategically if we assume that u_2 is uniformly distributed between u_3 and u_1 (that is, when $\tilde{u} \sim U[0,1]$).

Finally, Figure 4f illustrates the effects of marginal increases in the distance from contention and winning margin. (Once again, these effects are averaged across the 106 constituencies.) It reinforces our formal results: the incentive is increasing in both the distance from contention (a standard prediction) and the expected winning margin (contrary to the marginality hypothesis) but the effect of the winning margin is much weaker than the distance from contention.

In summary, this exercise suggests that the distance from contention should be a clear driver of strategic-voting incentives; it suggests that the effect of the winning margin should be positive but more muted (for example, the distance from contention is between three and seven times more powerful across the range of s considered here); and it suggests that a substantial fraction of third-party supporters (over half if $s \geq 16$ and $\tilde{u} \sim [0,1]$) should vote strategically.

[21] In calculating the "average strategic incentive" we have weighted by the number of Liberal Democrat voters in each constituency; however, very similar results hold for other weighting schemes.

Individual-Level Data

The British Election Studies (BES) provide high-quality post-election survey data with sufficient questions to study strategic voting in depth (Heath, Jowell, and Curtice, 1995; Heath, Jowell, Curtice, Brand, and Mitchell, 1996; Heath, Jowell, Curtice, and Norris, 2000; Clarke, Sanders, Stewart, and Whiteley, 2003, 2006, 2011). Six surveys from 1987 to 2010 offer comparable high quality face-to-face random probability samples that reveal strategic motivations for vote choice.

Measuring Strategic Voting. There is substantial debate about how to measure strategic voting (Alvarez, Boehmke, and Nagler, 2006; Blais, Young, and Turcotte, 2005; Fisher, 2004; Herrmann and Pappi, 2008; Herrmann, 2010). To evaluate both the intuitive predictions and the theory developed in this paper with individual-level data we need strict preference orderings for the three main parties and a 'direct' measure so that strategic voters can be identified without employing a theory of strategic voting incentives as the 'indirect' methods do. We follow the method developed by Fisher (2004) which is based primarily on the following survey question.[22]

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).*

Strategic voters were mainly identified as those who gave response (3) which identifies the act of deserting the preferred party and links it to the cause (the perceived weakness of the preferred party). All strategic voters ought to choose option (3), but respondents who gave tactical reasons in option (4) were also coded as strategic. Respondents who gave any indication that they voted for their preferred party were not coded as strategic. This was done using a follow-up question asking for the "party you really preferred" for those who gave option (3), and using strength-of-feeling scores. These are derived from the following question from the 1987 to 1997 surveys.

[22] After 2000 responses (1) and (2) changed to "The party had the best policies" and "had the best leader." This does not seem to have materially affected the efficacy of the question for identifying strategic voters (Fisher, 2014). We recognize that (in both cases) the response set is odd and incomplete (Blais, Young, and Turcotte, 2005). However, (3) is sufficiently clear to make the other options irrelevant to those who have strategically deserted their preferred party.

Please choose a phrase from this card to say how you feel about the ← party name inserted here →

(1) *Strongly in favour*

(2) *In favour*

(3) *Neither in favour nor against*

(4) *Against*

(5) *Strongly against*

For the 2001 to 2010 surveys respondents were asked to indicate how they felt about each party on a 0 to 10 scale from "strongly dislike" to "strongly like". For pooled analysis both questions were rescaled to run from 0 (most hostile) to 1 (most positive) response.

When the respondent was a strategic voter the preferred party was provided either by the "party you really preferred" question or via the strength-of-feeling scores. For non-strategic voters the party voted for was the first-preference party, unless there was a clear indication otherwise on the strength-of-feeling scores. Second- and third-preference parties were identified similarly.

Both strength-of-feeling scores and information about how people would have used a second vote were used for consistency checks and to identify strict preference rankings (Fisher, 2004).[23]

Sample. We focus on voters in England. The major three parties took the top three positions in nearly all English constituencies and all of those in the BES surveys.

The observed rates of strategic voting are reported in Table 2. If a preferred party is running first or second then strategic voting incentives are (theoretically) either absent or muted, and so Table 2 also reports rates of strategic voting for restricted sub-samples: the theoretical risk population, where a voter's first preference trailed behind her second preference; the narrower population of third-party supporters; and another population consisting of those whose preferred party came third or lower in the current election, previous election, or in a uniform-change constituency prediction from the final week national opinion polls. Strategic voting is much more common among third-party supporters, but it is not ubiquitous.[24]

[23]For the remaining set of otherwise unbreakable ties between second and third preference, the allocation prejudiced against our own theory. In such cases, for a strategic voter the second choice was the party with the lowest vote share of the two. For non-strategic voters, the second preference was the party with the highest vote share of the two parties. This is a "worst case scenario" for our theory. Results under other coding schemes are similar.

[24]It is rare but not absent from other-party supporters. Because of the different baseline population sizes there are almost as many voters deserting competitive parties as there are classic strategic voters.

Election	All Voters %	N	Risk Population %	N	Third Party Supporters (i) %	N	Third Party Supporters (ii) %	N
1987	4.6	2453	11.4	795	13.9	460	12.3	578
1992	6.8	1851	21.2	473	27.5	300	21.6	406
1997	7.8	1648	21.5	498	28.2	334	25.9	410
2001	7.2	1061	17.2	335	21.4	210	22.4	236
2005	5.5	1402	12.2	485	16.5	318	15.1	375
2010	5.7	1286	13.8	456	20.3	238	18.2	288
Total	6.3	9701	15.9	3042	21.1	1860	19.1	2293

Notes. The data are from the British Election Studies. Case (i) of third-party supporters comprises those whose preferred party came their in their constituency. Case (ii) adds those whose party came third inn the previous election, or in a nationwide poll projection. We restrict to sincere and strategic voters, and to supporters of the three major parties.

TABLE 2. Rates of Strategic Voting

Likewise, there are non-strategic non-sincere voters: they voted for one major party but liked another more, and yet they did not report a strategic motivation. There are other motivations for such behavor, such as protest voting; such voters are excluded from our analysis. Similarly, we exclude those who preferred or voted for a minor party. Finally, the (very few) respondents who did not report all of the strength-of-feeling scores are also excluded.

Three observations emerge. Firstly, strategic voting is significant: for 1997 (when the Conservatives were removed from government) over a quarter of third-party supporters voted strategically. Secondly, this pattern is inconsistent with both Duvergerian and non-Duvergerian equilibria from classic theories. Thirdly, strategic voting seems slightly low in the light of Figure 4.

Expanding on this third point, and inspecting Figure 4e, note that for $s = 20$ (the precision obtained from a random sample of 20 voting intentions) the average strategic incentive faced by a supporter of a third-placed Liberal Democrat is $\Lambda \approx 0.56$. Given a uniformity assumption on $\tilde{u}$, this is approximately double the empirical incidence. This suggests that either voters' beliefs are rather less precise than this or, perhaps, that not all voters are instrumentally motivated.

Constituency Shares and Voters' Preferences. A voter's incentive to vote strategically theoretically depends upon the expected vote shares (Lemma 3). These are proxied by the actual vote shares. Given these proxies, and for the analysis of third-party supporters, *distance from contention* is the difference in between the preferred party and the second-place party; the *margin of victory* is the difference between winning and second-placed parties.

Measurement of the relative strength of preference is based on the strength-of-feeling scores rescaled to $[0, 1]$. The score for the favorite minus that for the second choice is the *first preference gap*; the *second preference gap* is defined similarly. All theories suggest that strategic voting should be decreasing in the first preference gap and increasing in the second. Given the absence of proper measures of von Neumann-Morgenstern utilities, these gaps could be considered as either useful proxies or relevant covariates in the evaluation of our theoretical predictions.

Finally, we sometimes use election dummies as covariates. For models without preference gaps as covariates the coefficients of the election dummies will in part reflect temporal variation in the average of the relative preference component of the voting decision rule (Proposition 1). They also allow for untheorized variation in the propensity to vote strategically over time.

Individual-level Regressions

Our main theoretical results relate the incentive to vote strategically to a voter's perception of the constituency situation (Propositions 5–6). Here we investigate this empirically: Table 3 reports two OLS and two probit specifications for strategic voting amongst third-party supporters.

Distance from Contention and Marginality. There is a positive association between strategic voting and a third-placed party's distance from contention. For example, pushing that distance up by 5% corresponds to a 4% increase in the probability of a strategic vote; this effect emerges from the OLS estimates of a linear probability model. Under a probit specification the average marginal effect of the distance from contention (from the second probit model) ranges from 0.5 (a 5% increase in distance raises strategic voting by 2.5%) when the preferred party is very close to second place, to 1.0 when the preferred party is 35 points behind. These figures are calculated holding the margin and preference gaps at their mean and averaging over elections.

Finding 1 (Distance from Contention)**.** *Controlling for the size of the margin of victory, strategic voting increases with the distance from contention of a preferred third-placed candidate.*

Traditional intuition and our formal results disagree on the relationship between the closeness of a race and the incentive to vote strategically: our formal result suggests that the effect of a wider margin should be positive rather than negative, and it should be relatively weak.

From Table 3, there is essentially no evidence for a partial association between the margin of victory and strategic voting after controlling for the distance from contention. The results are

	OLS (i)	OLS (ii)	Probit (i)	Probit (ii)
Constant	-0.00	0.03	⋆ -1.61	⋆ -1.50
	(0.04)	(0.04)	(0.14)	(0.17)
Distance from Contention	⋆ 0.87	⋆ 0.80	⋆ 3.07	⋆ 2.93
	(0.13)	(0.12)	(0.47)	(0.50)
Margin of Victory Actual	0.15	0.11	0.60	0.34
	(0.10)	(0.10)	(0.36)	(0.38)
1st Preference Gap (1st-2nd)		⋆ -0.45		⋆ -2.20
		(0.05)		(0.23)
2nd Preference Gap (2nd-3rd)		⋆ 0.30		⋆ 1.08
		(0.04)		(0.16)
Election (1987 baseline)				
1992	⋆ 0.10	‡ 0.10	⋆ 0.36	‡ 0.37
	(0.03)	(0.03)	(0.11)	(0.12)
1997	⋆ 0.12	‡ 0.09	⋆ 0.45	‡ 0.36
	(0.03)	(0.03)	(0.10)	(0.11)
2001	0.06	† 0.07	† 0.24	† 0.31
	(0.03)	(0.03)	(0.12)	(0.13)
2005	0.04	0.05	0.16	0.21
	(0.03)	(0.03)	(0.11)	(0.12)
2010	0.06	† 0.07	† 0.26	† 0.32
	(0.03)	(0.03)	(0.12)	(0.13)
BIC	1946.26	1770.69	1914.05	1728.05

Notes. We use $N = 1860$ observations from the British Election Studies: these are sincere and strategic supporters of third-placed parties in England. Each entry reports the estimated coefficient, with the standard error in parentheses below. Statistical significance at the 0.05 (†), 0.01 (‡), and 0.001 ($\star$) levels is indicated with the corresponding symbol.

TABLE 3. Regression Models of Strategic Voting for Third-Party Supporters

closest to our formal model which suggests there should be a small positive relationship. The primary reason that our estimated coefficients are not statistically significant at conventional levels is that there are not enough observations to provide a sufficiently powerful test which distinguishes between our theory and the null hypothesis of no effect.

Finding 2 (Margin of Victory)**.** *Controlling for the distance from contention of a preferred third-placed party, strategic voting increases, very weakly, with the size of the margin of victory between the first-placed and second-placed parties.*

The coefficients on the distance from contention and margin of victory may be compared to the theoretical coefficients from our Dirichlet model. Inspecting Figure 4f, for low levels of belief precision we see a theoretical effect of an increase in d which is twice as high as its empirical counterpart; for moderate and higher levels the theoretical effect is four times as high. One explanation for this is the usual attenuation of coefficients from the noisy measurement of covariates. In this case, the ideal covariates would be the electoral situation perceived by the voter; however, we proxy for this via the actual election outcome. A second explanation revisits an earlier theme: we suggest that not all voters pursue short-term instrumental objectives when selecting their votes. For example, if only a fraction are instrumental then the reaction of strategic voting to constituency parameters (such as d and m) scales down accordingly.

Other Specifications. The second and fourth models in Table 3 include the preference gaps between the preferred and second choice party and between the second and third choice; these are calculated from the strength-of-feeling scores. The coefficients confirm that third-party supporters are more likely to make a strategic switch if they are relatively indifferent between the two parties they like the most (as indicated by the negative coefficient for the first preference gap) and have a strong preference for their second over their third choice (as shown by the positive coefficient for the second preference gap). This is as expected from both intuition and from formal theory. The election dummies pick up some temporal variation unexplained by the theory.

These conclusions are robust to other specifications including the use of constituency random effects, weighting, interaction with election effects, more detailed parameterization allowing strength-of-feeling score effects to take arbitrary functional forms, use of three-party rather than actual constituency vote shares, and using the broader risk population from Table 2 instead of current third-party supporters. There was no evidence for any variation in the effects of constituency characteristics over time and some minor but statistically significant idiosyncratic variation in the strength of preference gap effects. Results are similar but unsurprisingly weaker if previous election results are used instead of current ones.

Overall the findings from this analysis accord more closely to the theory expounded here than to traditional intuition. While both theories agree about the effect of the distance from contention and voters' preferences, and the empirical analysis supports them, the effect of the margin of victory is not clearly negative as intuition suggests. The estimated coefficient, albeit statistically insignificant, is closer to the small positive effect predicted by our formal theory.

Calibrating the Voter Uncertainty Model

Our empirical coefficients lend some support to our comparative-static predictions. Here we take a small step towards a full structural specification by calibrating our voter uncertainty model.

The Average Strategic Incentive. Consider a voter from the BES sample with preferences $u_1 > u_2 > u_3$. We use the Dirichlet specification, $f(p) \propto (p_1^{\pi_1} p_2^{\pi_2} p_3^{\pi_3})^s$, and we proxy for π with the actual three-party vote shares. We ask: how strong will be her incentive to vote strategically? Recall (Proposition 1) that she should vote strategically if and only if $\tilde{u} < \Lambda$ where $\tilde{u} = (u_1 - u_2)/(u_1 - u_3)$ is the strength of her first preference and where Λ is from Lemma 3.[25] Finally, recall again that if $\tilde{u} \sim U[0,1]$, so that a voter's second choice is equally likely to be anywhere between her first and third choice, then Λ is the probability that she optimally votes strategically.

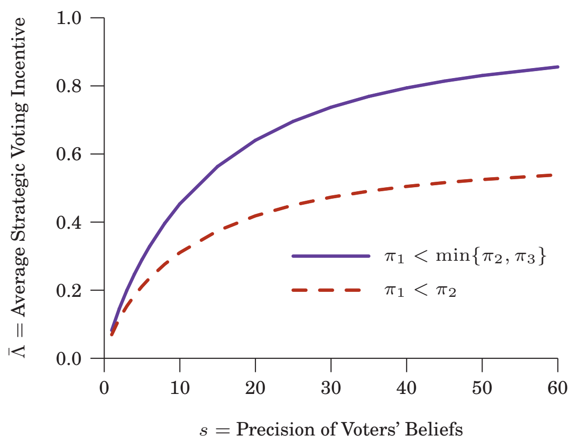
For various values of s, we calculated Λ for each member of the BES samples for whom the incentive was positive ($\Lambda > 0$, or when $\pi_1 < \pi_2$) and also for the sub-sample of third-party supporters (so that $\pi_1 < \min\{\pi_2, \pi_3\}$). Looking across our BES samples, we then calculated the average strategic incentive $\bar{\Lambda}$. These averages, plotted against s, are illustrated in Figure 5a.

If $\tilde{u}$ is distributed uniformly, then the proportion of instrumental voters who act strategically should be $\bar{\Lambda}$. Inspecting Figure 5a, if $s = 10$ (corresponding to a random sample of approximately ten others) then $\bar{\Lambda} = 0.4533$ across third-party supporters; double the empirical rate (Table 2).

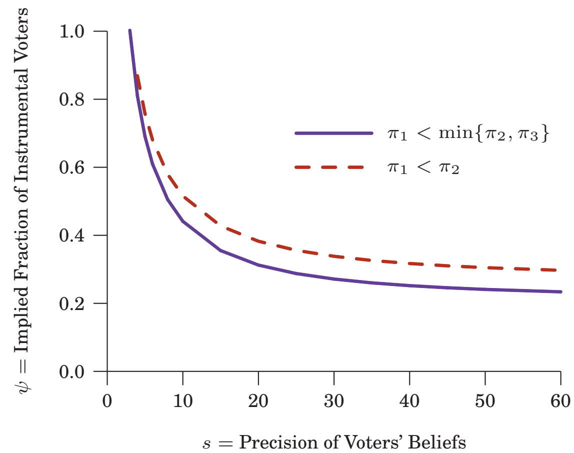
We highlight three ways in which the voter uncertainty model can match the empirical frequency of strategic voting. Firstly, the precision of voters' beliefs might be extremely low: to match an incidence of 20% amongst third-party supporters requires $s \approx 3$. Secondly, the distribution of voter types could deviate from $\tilde{u} \sim U[0,1]$: if voters care relatively little for their second preference then $\tilde{u}$ is large, which reduces $\Pr[\tilde{u} < \Lambda]$. Thirdly, not all voters are (narrowly) instrumental: their votes are motivated by concerns other than the identity of the election winner.

To explore the balance between the first and third features mentioned above, we fix the distribution of voter preferences so that $\tilde{u} \sim U[0,1]$. If a fraction ψ of voters have instrumental objectives then for voters with modal beliefs π the proportion who vote strategically should be $\psi \times \bar{\Lambda}$. Equating this to the observed frequency of strategic voters in our sample pins down an inverse relationship between ψ and the precision s of voters' beliefs.

[25] Recall that, explicitly, $\Lambda = [G(\pi_1, s) - G(\pi_2, s)]/[G(\pi_1, s) + 2G(\pi_3, s)]$ where $G(\pi_i, s) \propto \int_{1/3}^{1/2} [z^{1-\pi_i}(1-2z)^{\pi_i}]^s \, dz$. With a change of variable, $G(\pi_i, s) \propto 2^{\pi_i s} \int_0^{1/3} [z_i^{\pi}(1-z)^{1-\pi_i}]^s \, dz = 2^{\pi_i s} \, \mathrm{Beta}(1/3; \pi_i s + 1, (1-\pi_i)s + 1)$, where $\mathrm{Beta}(t; a, b)$ is the incomplete beta function evaluated at t with parameters a and b; this is readily calculated.



(A) Average Strategic Incentive



(B) Implied Fraction of Instrumental Voters

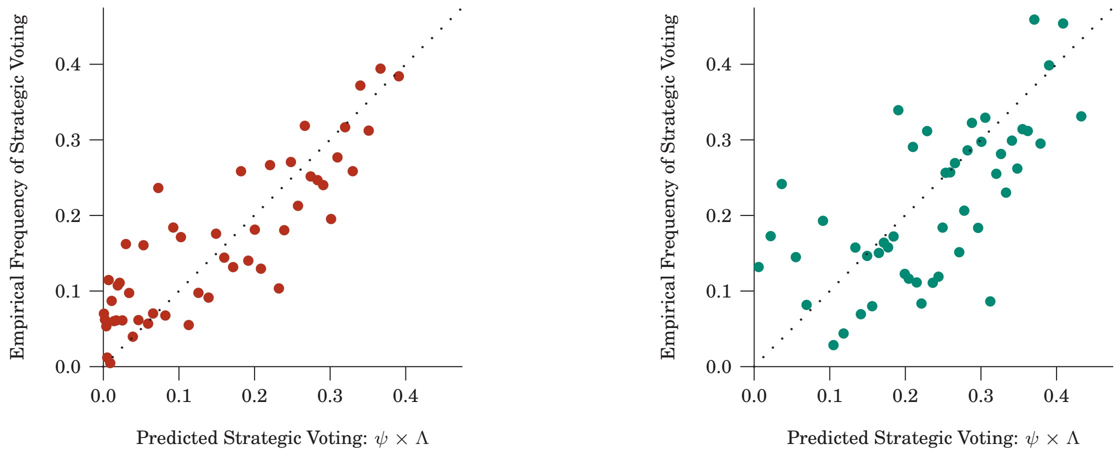
Notes. Panel (a) reports the average strategic incentive, whereas panel (b) reports the fraction of instrumental voters that is consistent with observed strategic voting amongst (i) the theoretical risk population (≈16%) and (ii) amongst third-party supporters (≈20%).

FIGURE 5. Belief Precision and Strategic Incentives in English Constituencies

The solid line in Figure 5b corresponds to the narrower third-party-supporters risk population, amongst whom approximately 20% voted strategically. Each s yields an average value of the strategic incentive $\bar{\Lambda}$. The vertical axis reports the implied fraction ψ of instrumental voters: this satisfies $\psi \times \bar{\Lambda} = 0.2$. (The broken line repeats the exercise using the (larger) theoretical risk population for whom $\pi_2 > \pi_1$.) The empirical frequency of strategic voting is consistent with a world in which a small fraction of voters are instrumental, but have good information; or with beliefs that are imprecise but where more voters pursue a short-term instrumental goal.[26]

This calibration exercise that lies behind Figure 5b can match prediction and observation at a disaggregated level. To illustrate this, we fixed $s = 10$ and calculated (as before) the predicted strategic incentive Λ for each BES sample member. Setting $\tilde{u} \sim U[0,1]$, we then calculated the probability of a strategic vote: $\psi \times \Pr[\tilde{u} < \Lambda] = \psi \times \Lambda$ where ψ is the proportion of voters who are instrumental. We set $\psi = 0.441$, so that the aggregate theoretical strategic voting frequency matches the empirical frequency. We then ordered the BES sampled voters by the predicted strategic voting incentive, grouped them into 50 ordered groups, and then calculated the average predicted strategic voting probability and the empirical strategic voting frequency for each group. The output from this exercise is illustrated as a simple scatter plot (Figure 6a). We performed the same exercise using the narrower risk population of third-party supporters (Figure 6b).

[26]This exercise is a complement to other recent work (Spenkuch, 2017) which seeks to pin down voters' motivations.



(A) Theoretical Risk Population ($\pi_1 < \pi_2$) (B) Third Party Supporters ($\pi_1 < \min\{\pi_2, \pi_3\}$)

Notes. These figures illustrate the correlation between theoretically predicted and empirically observed rates of strategic voting. We set $s = 10$, $\tilde{u} \sim U[0, 1]$, and $\psi = 0.44$.

FIGURE 6. Predicted versus Empirical Strategic Voting Rates

This rough-and-ready calibration yields a nice fit of theory and observation. The anomaly is for low values of the predicted strategic incentive: the empirical frequency is too high. These are constituencies where a voter's first and second choices are close, and so the strategic incentive should be small. However, here some voters may believe that they face a strong strategic incentive, even though they would not if their modal beliefs were closer to the election outcome.

Calibration via Maximum Likelihood. Recall that $\tilde{u}$ measures a voter's allegiance to her true first preference. We do not observe this directly and so (as before) we specify $\tilde{u} \sim U[0, 1]$. Fixing her (Dirichlet) beliefs, the probability that she acts strategically is Λ. Here, however, we acknowledge again that only a subset of voters (with probability ψ) are instrumental. Hence:

$$\Pr[\text{Strategic Vote}] = \psi \times \Lambda \text{ where } \Lambda = \frac{G(\pi_1, s) - G(\pi_2, s)}{G(\pi_1, s) + 2G(\pi_3, s)}. \tag{18}$$

Given that we do not see the preference type $\tilde{u}$ of a voter, her decision is uncertain. The other parameters (π, s, and ψ) are also unknown to us. To proceed, we proxy for π with the three-party vote shares. We then view s (belief precision) and ψ (the proportion of instrumental voters) as estimable parameters. We proceed to calibrate our model via maximum likelihood. In essence, this exercise picks one point on the instrumentality-information trade-off line from Figure 5b.

Proceeding in this way brings some challenges. The predicted probability of a strategic vote in equation (18) falls to zero as π_1 and π_2 become close. Nevertheless, there are observed situations in which a voter's first-preference and second-preference candidates run close and yet she votes

	(1)	(2)	(3)	(4)	(5)	(6)
	$d > 0.000$	$d > 0.010$	$d > 0.025$	$d > 0.000$	$d > 0.010$	$d > 0.025$
ψ	0.32 ⋆	0.38 ⋆	0.41 ⋆	0.32 ⋆	0.36 ⋆	0.40 ⋆
	(0.040)	(0.055)	(0.066)	(0.066)	(0.071)	(0.072)
s	21.4 ⋆	14.5 ⋆	11.9 ⋆	9.34 ‡	9.38 ‡	10.4 ‡
	(6.47)	(4.11)	(3.35)	(2.91)	(2.92)	(3.28)
ε				0.076 ⋆	0.057 ‡	0.023
				(0.017)	(0.020)	(0.025)
$1 - \varepsilon - \psi$				0.60 ⋆	0.58 ⋆	0.58 ⋆
				(0.062)	(0.068)	(0.070)
N	1860	1771	1667	1860	1771	1667
BIC	1934.8	1832.6	1735.2	1904.1	1829.1	1741.7

Notes. Statistical significance at the 0.05 (†), 0.01 (‡), and 0.001 (⋆) levels is indicated.

TABLE 4. Maximum Likelihood Estimates

strategically. This is inconsistent with equation (18). As we show here, this can cause difficulties for maximum-likelihood estimation across the full BES sample.[27] Furthermore, the driving force behind this is that voters' beliefs about their electoral situation are not always well proxied by the election outcome: there is variation in voters' information, which leads some of them to vote strategically even though (given the true electoral situation) it makes limited sense to do so.[28]

Nevertheless, we can proceed to obtained maximum-likelihood estimates both for the fuller samples and also for mildly restricted samples which exclude those cases where a voter's preferred candidate is in close contention. Our results are reported as specifications (1) to (3) in Table 4. All of these specifications restrict to cases where the voter's preferred candidate was in third place.

Specification (3) restricts to cases where the distance from contention exceeded 2.5%. In these cases, the strategic incentive is non-negligible. The estimates suggest that the precision of voters' beliefs correspond to $s \approx 12$, and that approximately 40% of voters are instrumentally motivated. In contrast, specifications (1) and (2) include voters who face a much smaller distance from contention. For these cases (including over 100 cases, for specification (1), where that distance fell below 1%) the theoretical strategic incentive is, for lower s, very small. The maximum-likelihood fit needs to employ larger values of s in order to accommodate strategic voting in such cases. In

[27] The estimator searches for the pair ψ and s which maximizes the likelihood function. Strategic voters for whom π_1 and π_2 are close penalize that likelihood function heavily; indeed, if $\pi_1 > \pi_2$ then the likelihood function says that a strategic vote is impossible. The maximization reacts to this by pursuing very large values of s (so that the probability of a strategic vote can be high whenever π_1 and π_2 are close). To fit the aggregate incidence of strategic voting then requires ψ to fall. The general pattern is that s diverges (to a high value) in the search for an estimate.

[28] A fuller solution would be to model fully the heterogeneity of voters' beliefs. This, however, is some distance beyond the scope of this paper and is a natural part of our ongoing research project.

turn, higher values of s need to be paired with lower values of ψ in order to fit the aggregate level of strategic voting. Thus, including all third-party supporters results in very different estimates that are largely driven by observations where $d \approx 0$. These are cases where voters' unobserved modal beliefs about their situation are likely to differ from the actual election outcome.

Equation (18) does not allow for voters who (given π) should not vote strategically (hence $\Lambda \approx 0$) but nevertheless do so. To allow space for them we introduce a noise term. We suppose that for some ε, a fraction ε always vote strategically; a further fraction ψ vote according to the Dirichlet model; and the remaining fraction $1 - \varepsilon - \psi$ are non-instrumental and so vote sincerely. Thus:

$$\Pr[\text{Strategic Vote}] = \varepsilon + (\psi \times \Lambda) \text{ where } \Lambda = \frac{G(\pi_1, s) - G(\pi_2, s)}{G(\pi_1, s) + 2G(\pi_3, s)}. \tag{19}$$

Allowing $\varepsilon > 0$ gives the model room to explain strategic voting even when the distance from contention is small. We proceed to estimate jointly ε, Λ, and ψ via maximum likelihood: see specifications (4)–(6) of Table 4. Very naturally, including lower values of the distance from contention pushes up the estimates of ε; this is needed to accommodate those who vote strategically when the preferred candidate is very much in contention. All three specifications generate similar estimates for s. The estimate for $\varepsilon + \psi$, which measures the maximum possible extent of strategic voting, is also approximately constant across these three specifications. Our results from this final calibration exercise are summarized in our third and final empirical finding.

Finding 3 (Calibration)**.** *The pattern of strategic voting across England matches a specification where a little over one third of voters are instrumentally motivated, where there is non-negligible uncertainty over likely vote shares, and where the precision of voters' beliefs corresponds to that obtained when a voter observes the voting intentions of ten to twelve other randomly selected voters.*

Concluding Remarks

A striking criticism of rational-choice theory is that it merely formalizes intuitive ideas (Green and Shapiro, 1994); but here, our predictions differ from intuition. Also lacking from many established theories of strategic voting are comparative-static predictions. Here we have offered such predictions in the presence of aggregate uncertainty.[29] We have demonstrated that the marginality effect should be both weak and opposite to that suggested by traditional intuition.

[29] In recent years models with aggregate uncertainty have been used to study strategic voting (Myatt, 2007; Fisher and Myatt, 2002b), the coordination of party members (Dewan and Myatt, 2007), voter turnout (Myatt, 2015; Evren, 2012), protest voting (Myatt, 2016), and measures of voter power (Le Breton, Lepelley, and Smaoui, 2014).

Furthermore, our game-theoretic model incorporates non-strategic types: when such sincere voters are present, we find an increase in the willingness of instrumental voters to act strategically.

We have evaluated our predictions in the context of British Election Study data. Our theory and traditional intuition offer the same predictions concerning the distance from contention. However, the traditional marginality hypothesis suggests that strategic voting should fall as the margin of victory widens. The data do not agree, and instead support our own prediction.[30]

Finally, we have calibrated a voter uncertainty model. That exercise disentangles the two reasons (the presence of sincere voters, or poor belief precision) that strategic voting is not more prevalent. A rough summary is that a little more than a third of voters are instrumental, with information equivalent to a modest sample of around ten to twelve independent voting intentions.

Omitted Proofs

Proof of Lemma 1. This is a special case of Lemma 1 from Myatt (2015) which concerns an election with m candidates and so generalizes classic results by Good and Mayer (1975) and Chamberlain and Rothschild (1981). For completeness, we prove here the three-candidate case. Using the complete expression for $\Pr[x]$,

$$\begin{aligned}\frac{(n+2)!}{n!}\Pr[b] &= \int_\Delta \frac{(n+2)!}{x_1!x_2!x_3!}\left[\prod\nolimits_{i=1}^3 p_i^{x_i}\right] f(p)\,dp \\ &= f\left(\frac{x}{n}\right)\underbrace{\int_\Delta \frac{(n+2)!}{x_1!x_2!x_3!}\left[\prod\nolimits_{i=1}^3 p_i^{x_i}\right] dp}_{\text{Dirichlet density}} + \int_\Delta \frac{(n+2)!}{x_1!x_2!x_3!}\left[\prod\nolimits_{i=1}^3 p_i^{x_i}\right]\left[f(p)-f\left(\frac{x}{n}\right)\right] dp \\ &= f\left(\frac{x}{n}\right) + \underbrace{\int_\Delta \frac{(n+2)!}{x_1!x_2!x_3!}\left[\prod\nolimits_{i=1}^3 p_i^{b_i}\right]\left[f(p)-f\left(\frac{x}{n}\right)\right] dp}_{\text{vanishes uniformly}}.\end{aligned}$$

We find a uniform bound for the final term, so showing that it declines uniformly to zero.[31] For $\gamma \in \Delta$ define $\Delta_\varepsilon^\gamma = \{p \in \Delta \mid \max|p_i - \gamma_i| \le \varepsilon\}$ as an ε neighborhood of γ. $f(\cdot)$ is a continuous density with bounded derivatives and so there is some positive D, which does not depend on γ, such that $|f(p) - f(\gamma)| \le D\varepsilon$ for all $p \in \Delta_\varepsilon^\gamma$. Furthermore, $\bar{f} = \max_{p\in\Delta} f(p)$ bounds the difference $|f(p) - f(\gamma)|$ for all $p \notin \Delta_\varepsilon^\gamma$. Hence

$$\begin{aligned}\left|\frac{(n+2)!}{n!}\Pr[x] - f\left(\frac{x}{n}\right)\right| &\le \int_\Delta \frac{(n+2)!}{x_1!x_2!x_3!}\left[\prod\nolimits_{i=1}^3 p_i^{x_i}\right]\left|f(p)-f\left(\frac{x}{n}\right)\right| dp \\ &\le D\varepsilon + \bar{f}\int_{\Delta/\Delta_\varepsilon^{x/n}} \frac{(n+2)!}{x_1!x_2!x_3!}\left[\prod\nolimits_{i=1}^3 p_i^{x_i}\right] dp = D\varepsilon + \bar{f}\int_{\Delta/\Delta_\varepsilon^{x/n}} \frac{(n+2)!}{n!}\Pr[x\mid p]\,dp.\end{aligned}$$

[30] The marginality hypothesis has found support elsewhere. Blais and Nadeau (1996) reported evidence in the context of Canadian elections. This could provide evidence that voters are not narrowly instrumental. Whereas the absolute probability of a pivotal outcome (which is higher in a closely fought marginal constituency) is not directly relevant to a voter who cares only about the winning candidate, broader objectives resurrect the importance of such probabilities. For example, if a voter has expressive concerns, or faces a personal ethical cost for shifting away from her true favorite, then the absolutely probability of a tie matters once more. Arguably, the effect of the winning margin on strategic voting provides evidence of whether narrow or broad motives lie behind the decisions of instrumentally rational voters.

[31] This proof uses a shorter version of the approach used in the supplement to Hummel (2012).

For $p \notin \Delta_\varepsilon^{x/n}$ there is some i such that $p_i < (x_i/n) - \varepsilon$ or $p_i > (x_i/n) + \varepsilon$. Consider the latter case (the former case is similar) and note that x_i is a binomial with parameters p_i and n. Hence

$$\Pr[x \,|\, p] \leq \Pr\left[\frac{x_i}{n} \leq p_i - \varepsilon\right] \leq e^{-2n\varepsilon^2},$$

where the second inequality is well known (Hoeffding, 1963). It follows that

$$\left|\frac{(n+2)!}{n!}\Pr[x] - f\left(\frac{x}{n}\right)\right| \leq D\varepsilon + \bar{f}\int_{\Delta/\Delta_\varepsilon^{x/n}} \frac{(n+2)!}{n!}\Pr[x\,|\,p]\,dp$$

$$\leq D\varepsilon + \bar{f}e^{-2n\varepsilon^2}\frac{(n+2)!}{n!}\int_{\Delta/\Delta_\varepsilon^{x/n}} dp \leq D\varepsilon + \bar{f}e^{-2n\varepsilon^2}\frac{(n+2)^2}{2}.$$

This bound holds for any ε. Take, for example, $\varepsilon = n^{-1/4}$. For this choice of ε

$$\left|\frac{(n+2)!}{n!}\Pr[x] - f\left(\frac{x}{n}\right)\right| \leq \frac{D}{n^{1/4}}\varepsilon + \bar{f}e^{-2\sqrt{n}}\frac{(n+2)^2}{2}.$$

Noting that the exponential term dominates the polynomial term, this bound vanishes as $n \to 0$. □

Proof of Lemma 2. Consider an exact two-way tie between 1 and 2; other cases are similar. Note that $n\Pr[x_1 = x_2 > x_3] = \frac{1}{n}\sum_{y=\lceil n/3\rceil}^{\lfloor n/2\rfloor} n^2 \Pr[x_1 = x_2 = y]$. Using Lemma 1, each element converges uniformly:

$$\lim_{n\to\infty}\max_{y\in\{\lceil n/3\rceil,\ldots,\lfloor n/2\rfloor\}}\left|n^2\Pr[x_1 = x_2 = y] - f\left(\tfrac{y}{n}, \tfrac{y}{n}, \tfrac{n-2y}{n}\right)\right| = 0,$$

$$\Rightarrow \quad \lim_{n\to\infty} n\Pr[x_1 = x_2 > x_3] = \lim_{n\to\infty}\sum_{y=\lceil n/3\rceil}^{\lfloor n/2\rfloor}\tfrac{1}{n}f\left(\tfrac{y}{n}, \tfrac{y}{n}, \tfrac{n-2y}{n}\right). \tag{20}$$

The right-hand side defines a Riemann integral, and converges to $\int_{1/3}^{1/2} f(z, z, 1-2z)\,dz$ as $n \to \infty$. □

Proof of Proposition 1. This is obtained by assembling the various elements of Table 1, re-arranging, and replacing the probabilities of ties and near ties by terms p_{ij}. □

Proof of Lemma 3. Using the formula for the kernel of a Dirichlet density from equation (5),

$$p_{12} = \int_{1/3}^{1/2} f(z, z, 1-2z)\,dz \propto \int_{1/3}^{1/2}[z^{1-\pi_3}(1-2z)^{\pi_3}]^s\,dz = G(\pi_3, s). \tag{21}$$

Similar operations for p_{13} and p_{23} yield equations (6) and (7). Now,

$$\frac{\partial\log[z^{1-\pi}(1-2z)^\pi]}{\partial\pi} = \log\left[\frac{1-2z}{z}\right] < 0 \quad \text{for} \quad \frac{1}{3} < z < \frac{1}{2} \quad \Rightarrow \quad \frac{G(\varpi, s)}{\partial\varpi} < 0. \tag{22}$$

Finally, make the change of variable $w = 1 - 2z$ to obtain:

$$\int_{1/3}^{1/2}[z^{1-\varpi}(1-2z)^\varpi]^s\,dz = \frac{\int_0^{1/3}[w^\varpi(1-w)^{1-\varpi}]^s\,dw}{2^{(1-\varpi)s+1}} = \frac{2^{\varpi s}B(\frac{1}{3}, \varpi s+1, (1-\varpi)s+1)}{2^{s+1}} \tag{23}$$

where $B(t, a, b)$ is the incomplete beta function. □

Proof of Proposition 2. As noted in the main text, this follows directly from Lemma 3. □

Our results use properties of the function $G(\varpi, s)$. Lemma 5 identifies useful properties of this function.

Lemma 5. $G(\varpi, s) \equiv \int_{1/3}^{1/2}[z^{1-\varpi}(1-2z)^\varpi]^s\,dz$ *satisfies the following properties. Firstly,* $\log G(\varpi, s)$ *is a convex function of* ϖ*. Secondly,* $G(\varpi, s)$ *is a convex function of* ϖ*. Thirdly,*

$$\lim_{s\to\infty}\frac{\log G(\varpi, s)}{s} = \log g^\star(\varpi) \quad \textit{where} \quad g^\star(\varpi) \equiv \max_{z\in[1/3,1/2]} z^{1-\varpi}(1-2z)^\varpi. \tag{24}$$

Fourthly, $z^\star(\varpi) \equiv \arg\max_{z\in[1/3,1/2]} z^{1-\varpi}(1-2z)^\varpi = \max\{1/3,(1-\varpi)/2\}$*, and so* $g^\star(\varpi)$ *is strictly decreasing in* ϖ *for* $\varpi < \frac{1}{3}$*. Finally, if* $\pi_i < \min\{1/3, \pi_j\}$ *then* $\lim_{s\to\infty}[G(\pi_i,s)/G(\pi_j,s)] = \infty$*.*

Proof of Lemma 5. Firstly, we show that $\log G(\varpi,s)$ is a convex function of ϖ. Differentiating,

$$-\frac{\partial \log G(\varpi,s)}{\partial\varpi} = s\int_{1/3}^{1/2} \frac{[z^\varpi(1-2z)^{1-\varpi}]^s}{G(\varpi,s)} \log\left[\frac{z}{1-2z}\right] dz. \tag{25}$$

We show that the right-hand side of equation (25) is decreasing in π. Note that $[z^\varpi(1-2z)^{1-\varpi}]^s/G(\varpi,s)$ is a density over the interval $[1/3,1/2]$, and so the integral in equation (25) is the expectation of an increasing function. It is sufficient to show that a reduction in ϖ induces a first order stochastically dominant shift upward in the density $[z^\varpi(1-2z)^{1-\varpi}]^s/G(\varpi,s)$. We show that for any y where $1/3 < y < 1/2$:

$$\pi_j < \pi_k \quad \Rightarrow \quad \frac{\int_{1/3}^{y}[w^{\pi_j}(1-2w)^{1-\pi_j}]^s\,dw}{\int_{1/3}^{1/2}[z^{\pi_j}(1-2z)^{1-\pi_j}]^s\,dz} \le \frac{\int_{1/3}^{y}[w^{\pi_k}(1-2w)^{1-\pi_k}]^s\,dw}{\int_{1/3}^{1/2}[z^{\pi_k}(1-2z)^{1-\pi_k}]^s\,dz} \tag{26}$$

$$\Leftrightarrow \quad \frac{\int_{1/3}^{y}[w^{\pi_j}(1-2w)^{1-\pi_j}]^s\,dw}{\int_{1/3}^{y}[w^{\pi_k}(1-2w)^{1-\pi_k}]^s\,dw} \le \frac{\int_{1/3}^{1/2}[z^{\pi_j}(1-2z)^{1-\pi_j}]^s\,dz}{\int_{1/3}^{1/2}[z^{\pi_k}(1-2z)^{1-\pi_k}]^s\,dz}. \tag{27}$$

It is sufficient to show that the left hand side term is maximized at $y = 1/2$, and for this it is sufficient that its derivative (with respect to y) is positive on the interval $1/3 < y < 1/2$:

$$\frac{\partial}{\partial y}\log\left[\frac{\int_{1/3}^{y}[w^{\pi_j}(1-2w)^{1-\pi_j}]^s\,dw}{\int_{1/3}^{y}[w^{\pi_k}(1-2w)^{1-\pi_k}]^s\,dw}\right] \ge 0 \Leftrightarrow \frac{[y^{\pi_j}(1-2y)^{1-\pi_j}]^s}{\int_{1/3}^{y}[w^{\pi_j}(1-2w)^{1-\pi_j}]^s\,dw} \ge \frac{[y^{\pi_k}(1-2y)^{1-\pi_k}]^s}{\int_{1/3}^{y}[w^{\pi_k}(1-2w)^{1-\pi_k}]^s\,dw}. \tag{28}$$

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

$$\left[\frac{y(1-2w)}{w(1-2y)}\right]^{\pi_j} \le \left[\frac{y(1-2w)}{w(1-2y)}\right]^{\pi_k}. \tag{29}$$

This is true, since $\pi_j < \pi_k$ and $y(1-2w) \ge w(1-2y)$ for $1/3 < w \le y < 1/2$. We have demonstrated that the right hand side of equation (25) is decreasing in ϖ, as required.

Secondly, the convexity of $G(\varpi,s)$ follows directly: it is a convex function of a convex function.

Thirdly, we turn to the asymptotic behavior of the ratio of $\log G(\varpi,s)$ and s as $s\to\infty$. Both the numerator and denominator vanish to zero, and so we apply l'Hôpital's rule here:

$$\lim_{s\to\infty}\frac{\log G(\varpi,s)}{s} = \lim_{s\to\infty}\frac{\partial\log G(\varpi,s)}{\partial s} = \lim_{s\to\infty}\int_{1/3}^{1/2}\log[z^\varpi(1-2z)^{1-\varpi}]\left[\frac{[z^\varpi(1-2z)^{1-\varpi}]^s}{G(\varpi,s)}\right] dz. \tag{30}$$

The bracketed term in the right-hand side integrand is a density over $[1/3,1/2]$. As $s\to\infty$ it focuses all weight on $z^\star(\varpi)$; the proof of this follows the technique used to prove Lemmas 1 and 2.

Fourthly, $z^\star(\omega)$ is straightforwardly the solution to the maximization problem. If $\varpi < \frac{1}{3}$ then $z^\star(\varpi) = (1-\varpi)/2$ and $g^\star(\varpi)$ is strictly decreasing; if $\varpi \ge \frac{1}{2}$ then $g^\star(\varpi)$ is constant with respect to ϖ.

Finally, $\pi_i < \min\{1/3,\pi_j\}$ implies (using results so far) that $g^\star(\varpi)$ is strictly decreasing over the interval from π_i to π_j, and so $g^\star(\pi_i) > g^\star(\pi_j)$. Using the second claim of the lemma,

$$\begin{aligned}\lim_{s\to\infty}\frac{\log G(\pi_i,s)-\log G(\pi_j,s)}{s} &= \log g^\star(\pi_i) - \log g^\star(\pi_j) > 0 \\ \Rightarrow \quad \lim_{s\to\infty}\left[\frac{G(\pi_i,s)}{G(\pi_j,s)}\right] &= \lim_{s\to\infty}\exp\left(s\left[\frac{\log G(\pi_i,s)-\log G(\pi_j,s)}{s}\right]\right) = \infty. \quad \square\end{aligned}$$

Proof of Proposition 3. We begin with the case $\pi_2 > \pi_1 > \pi_3$. Using equation (7),

$$\Lambda = \frac{1 - [G(\pi_2, s)/G(\pi_1, s)]}{1 + 2[G(\pi_3, s)/G(\pi_1, s)]} < \frac{1}{1 + 2[G(\pi_3, s)/G(\pi_1, s)]}. \tag{31}$$

Candidate 3 is expected to place last, and so necessarily $\pi_3 < \min\{1/3, \pi_1\}$. Lemma 5 applies, and so $\lim_{s\to\infty}[G(\pi_3, s)/G(\pi_1, s)] = \infty$, and so $\lim_{s\to\infty} \Lambda = 0$ as claimed.

We now turn to the case where $\pi_1 < \min\{\pi_2, \pi_3\}$, which of course implies that $\pi_1 < \frac{1}{3}$. Lemma 5 applies again: $\lim_{s\to\infty}[G(\pi_2, s)/G(\pi_1, s)] = \lim_{s\to\infty}[G(\pi_3, s)/G(\pi_1, s)] = 0 \Rightarrow \lim_{s\to\infty} \Lambda = 1$. □

Proof of Proposition 4. By inspection of equation (7), Λ is increasing in π_2 and decreasing in π_1; it is increasing in π_3 so long as $G(\pi_1, s) > G(\pi_2, s)$, which holds because $\pi_1 < \pi_2$. □

Proof of Proposition 5. The effect of d is discussed in the text. We focus here on the margin of victory. We begin with $\pi_3 > \pi_2 > \pi_1$. Taking the solutions for π in terms of m and d from equation (8):

$$\frac{\partial \pi_1}{\partial m} = -\frac{1}{3}, \quad \frac{\partial \pi_2}{\partial m} = -\frac{1}{3} \quad \text{and} \quad \frac{\partial \pi_3}{\partial m} = \frac{2}{3}. \tag{32}$$

Using equation (7),

$$\Lambda = \frac{1 - [G(\pi_2, s)/G(\pi_1, s)]}{1 + 2[G(\pi_3, s)/G(\pi_1, s)]}. \tag{33}$$

$\pi_2 > \pi_1$ implies $G(\pi_2, s) < G(\pi_1, s)$, and so Λ is decreasing in $G(\pi_3, s)/G(\pi_1, s)$. From equation (32) an increase in m increases π_3 and lowers π_1; this reduces $G(\pi_3, s)$ and raises $G(\pi_1, s)$, and so Λ rises. To demonstrate the claim of the proposition we now differentiate and use equation (32) to obtain:

$$\frac{\partial \log[G(\pi_2, s)/G(\pi_1, s)]}{\partial m} = -\frac{1}{3}\left[\frac{\partial \log G(\pi_2, s)}{\partial \pi_2} - \frac{\partial \log G(\pi_1, s)}{\partial \pi_1}\right] < 0. \tag{34}$$

The inequality follows from the convexity of $\log G(\varpi, s)$ reported in Lemma 5 and the fact that $\pi_2 > \pi_1$. It follows that the numerator of Λ is increasing in m. Next, we consider the case $\pi_2 > \pi_3 > \pi_1$, so that

$$\frac{\partial \pi_1}{\partial m} = -\frac{1}{3}, \quad \frac{\partial \pi_2}{\partial m} = \frac{2}{3} \quad \text{and} \quad \frac{\partial \pi_3}{\partial m} = -\frac{1}{3}. \tag{35}$$

This time we write the incentive to vote strategically as

$$\Lambda = 1 - \frac{2 + [G(\pi_2, s)/G(\pi_3, s)]]}{2 + [G(\pi_1, s)/G(\pi_3, s)]}. \tag{36}$$

Clearly, $G(\pi_2, s)/G(\pi_3, s)$ is decreasing in m. Now, using the convexity of $\log G(\varpi, s)$ again,

$$\frac{\partial \log[G(\pi_1, s)/G(\pi_3, s)]}{\partial m} = \frac{1}{3}\left[\frac{\partial \log G(\pi_3, s)}{\partial \varpi} - \frac{\partial \log G(\pi_1, s)}{\partial \varpi}\right] > 0.$$ □

Proof of Proposition 6. We begin with $\pi_3 > \pi_2 > \pi_1$. Differencing the effect of d and m:

$$\frac{\partial \pi_1}{\partial d} - \frac{\partial \pi_1}{\partial m} = -\frac{1}{3}, \quad \frac{\partial \pi_2}{\partial d} - \frac{\partial \pi_2}{\partial m} = \frac{2}{3} \quad \text{and} \quad \frac{\partial \pi_3}{\partial d} - \frac{\partial \pi_3}{\partial m} = -\frac{1}{3}. \tag{37}$$

From this it follows that

$$\frac{\partial [G(\pi_2, s)/G(\pi_1, s)]}{\partial d} - \frac{\partial [G(\pi_2, s)/G(\pi_1, s)]}{\partial m} < 0, \tag{38}$$

and so the numerator of Λ in equation (33) is increasing in $d - m$. Turning to the denominator,

$$\frac{\partial \log[G(\pi_3, s)/G(\pi_1, s)]}{\partial d} - \frac{\partial \log[G(\pi_3, s)/G(\pi_1, s)]}{\partial m} = -\frac{1}{3}\left[\frac{\partial \log G(\pi_3, s)}{\partial \varpi} - \frac{\partial \log G(\pi_1, s)}{\partial \varpi}\right] < 0, \tag{39}$$

where the inequality comes once again from the convexity of $\log G(\varpi, s)$ in ϖ and $\pi_3 > \pi_1$. Combining these observations, this implies that Λ is increasing in $d - m$. For the case $\pi_2 > \pi_3 > \pi_1$,

$$\frac{\partial \pi_1}{\partial d} - \frac{\partial \pi_1}{\partial m} = -\frac{1}{3}, \quad \frac{\partial \pi_2}{\partial d} - \frac{\partial \pi_2}{\partial m} = -\frac{1}{3} \quad \text{and} \quad \frac{\partial \pi_3}{\partial d} - \frac{\partial \pi_3}{\partial m} = \frac{2}{3}. \tag{40}$$

Inspecting Λ in equation (33), it is straightforward to see that the denominator is decreasing in $d - m$. For the numerator, we again use the convexity of $\log G(\varpi, s)$:

$$\frac{\partial \log[G(\pi_2, s)/G(\pi_1, s)]}{\partial d} - \frac{\partial \log[G(\pi_2, s)/G(\pi_1, s)]}{\partial m} = -\frac{1}{3}\left[\frac{\partial \log G(\pi_2, s)}{\partial \varpi} - \frac{\partial \log G(\pi_1, s)}{\partial \varpi}\right] < 0, \tag{41}$$

and so the numerator is increasing in $d - m$.

Finally, we consider the effect of m when d is small. If $\pi_3 > \pi_2 > \pi_1$ the $d = \pi_2 - \pi_1$. Setting $d = 0$ yields $\Lambda = 0$; this does not change with m. The stated result then holds by continuity. □

Proof of Lemma 4 (for $\psi = 1$). Here, reference to voters is to the fraction $(1 - \pi_3)$ who dislike candidate 3 and so choose between candidates 1 and 2. Amongst those voters, and conditional on θ, we write v for the probability that a vote is cast for candidate 1: $v = \Pr[\log(u_1/u_2) + b\hat{\theta} \geq 0 \,|\, \theta]$. Now, note that

$$\log\left(\frac{u_1}{u_2}\right) + b\hat{\theta} = (\sigma + b)\theta + (\sigma + b)\eta + \sigma\varepsilon \sim N\left((\sigma + b)\theta, \frac{(\sigma + b)^2 + (\alpha - 1)\sigma^2}{\alpha}\right) \tag{42}$$

$$\Rightarrow \quad v = \Phi\left((\sigma + b)\theta\sqrt{\frac{\alpha}{(\sigma + b)^2 + (\alpha - 1)\sigma^2}}\right) \tag{43}$$

$$\Rightarrow \quad \theta = \frac{\Phi^{-1}(v)}{\sigma + b}\sqrt{\frac{(\sigma + b)^2 + (\alpha - 1)\sigma^2}{\alpha}} \tag{44}$$

$$\text{and} \quad \frac{dv}{d\theta} = (\sigma + b)\sqrt{\frac{\alpha}{(\sigma + b)^2 + (\alpha - 1)\sigma^2}}\,\phi\left((\sigma + b)\theta\sqrt{\frac{\alpha}{(\sigma + b)^2 + (\alpha - 1)\sigma^2}}\right) \tag{45}$$

$$= (\sigma + b)\sqrt{\frac{\alpha}{(\sigma + b)^2 + (\alpha - 1)\sigma^2}}\,\phi\left(\Phi^{-1}(v)\right), \tag{46}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution and density. Consider a voter with a signal $\hat{\theta}$, and write $g(\theta \,|\, \hat{\theta})$ and $h(v \,|\, \hat{\theta})$ for the densities of her posterior beliefs about θ and v. These densities are related: $h(v \,|\, \hat{\theta}) = g(\theta \,|\, \hat{\theta})/(dv/d\theta)$, where θ and v are related by equation (44). That is,

$$h(v \,|\, \hat{\theta}) = \frac{g\left(\theta \,|\, \hat{\theta}\right)\sqrt{\frac{(\sigma+b)^2+(\alpha-1)\sigma^2}{\alpha(\sigma+b)^2}}}{\phi\left(\Phi^{-1}(v)\right)} \quad \text{where} \quad \theta = \frac{\Phi^{-1}(v)}{\sigma + b}\sqrt{\frac{(\sigma + b)^2 + (\alpha - 1)\sigma^2}{\alpha}}. \tag{47}$$

Define $v^\dagger = \pi_3/(1-\pi_3) > \frac{1}{2}$ to be the coordination required to defeat candidate 3. Using the techniques used to prove Lemma 1, we can show that $\lim_{n\to\infty} n \Pr[x_1 = x_3 > x_2] = \lim_{n\to\infty} n \Pr[x_1 = x_3 - 1 > x_2] = h(v^\dagger \,|\hat{\theta})$, and similarly $\lim_{n\to\infty} n \Pr[x_2 = x_3 > x_1] = \lim_{n\to\infty} n \Pr[x_2 = x_3 - 1 > x_1] = h(1 - v^\dagger \,|\hat{\theta})$.

We note that $\Phi^{-1}(1 - v^\dagger) = -\Phi^{-1}(v^\dagger)$, and $\phi(\Phi^{-1}(1 - v^\dagger)) = \phi(\Phi^{-1}(v^\dagger))$. Hence,

$$\frac{h(v^\dagger \,|\hat{\theta})}{h(1 - v^\dagger \,|\hat{\theta})} = \frac{g(\theta^\dagger \,|\, \hat{\theta})}{g(-\theta^\dagger \,|\, \hat{\theta})} \quad \text{where} \quad \theta^\dagger = \frac{\Phi^{-1}(v^\dagger)}{\sigma + b}\sqrt{\frac{(\sigma + b)^2 + (\alpha - 1)\sigma^2}{\alpha}}. \tag{48}$$

Next, we turn to the density $g(\theta \,|\, \hat{\theta})$. Recall that $\theta \,|\, \hat{\theta} \sim N(\hat{\theta}, (1/\alpha))$, and so

$$g(\theta \,|\, \hat{\theta}) \propto \exp\left(-\frac{\alpha(\theta - \hat{\theta})^2}{2}\right) \quad \Rightarrow \quad \log\frac{g(\theta \,|\, \hat{\theta})}{g(-\theta \,|\, \hat{\theta})} = \frac{\alpha(\theta + \hat{\theta})^2 - \alpha(\theta - \hat{\theta})^2}{2} = 2\alpha\theta\hat{\theta}. \tag{49}$$

Putting all of our observations together,

$$\lim_{n\to\infty} \log \frac{\Pr[x_1 = x_3 > x_2 \mid \hat{\theta}]}{\Pr[x_2 = x_3 > x_1 \mid \hat{\theta}]} = \log \left[\frac{h(v^\dagger \mid \hat{\theta})}{h(1 - v^\dagger \mid \hat{\theta})} \right] = 2\hat{\theta}\Phi^{-1}(v^\dagger)\sqrt{\alpha + \frac{\alpha(\alpha-1)\sigma^2}{(\sigma+b)^2}}, \tag{50}$$

which yields the main claim of the lemma; the other claims are by inspection. □

Proof of Proposition 7. Existence and uniqueness are discussed in the text: specifically, $B(b)$ is decreasing and satisfies $B(0) > 0$ and so there is a unique solution $b^\star = B(b^\star)$. The various comparative-static claims follow directly from the response of the function $B(b)$ to changes in the parameters. □

Proof of Proposition 8. The comparative-static claims with respect to m and d are reported and proven by Myatt (2007). The additional result, regarding the negligible effect of marginality when the distance from contention is small, follows from the argument given in the main text. □

References

ABRAMSON, P. R., J. H. ALDRICH, P. PAOLINO, AND D. W. ROHDE (1992): "'Sophisticated Voting' in the 1988 Presidential Primaries," *American Political Science Review*, 86(1), 55–69.

ALVAREZ, R. M., F. J. BOEHMKE, AND J. NAGLER (2006): "Strategic Voting in British Elections," *Electoral Studies*, 25(1), 1–19.

BLAIS, A., AND R. NADEAU (1996): "Measuring Strategic Voting: A Two-Step Procedure," *Electoral Studies*, 15(1), 39–52.

BLAIS, A., R. YOUNG, AND M. TURCOTTE (2005): "Direct or Indirect? Assessing Two Approaches to the Measurement of Strategic Voting," *Electoral Studies*, 24(2), 163–176.

BOUTON, L., M. CASTANHEIRA, AND A. LLORENTE-SAGUER (2017): "Multicandidate Elections: Aggregate Uncertainty in the Laboratory," *Games and Economic Behavior*, 101, 132–150.

CAIN, B. E. (1978): "Strategic Voting in Britain," *American Journal of Political Science*, 22(3), 639–655.

CHAMBERLAIN, G., AND M. ROTHSCHILD (1981): "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory*, 25(1), 152–162.

CHHIBBER, P., AND K. KOLLMAN (1998): "Party Aggregation and the Number of Parties in India and the United States," *American Political Science Review*, 92(2), 329–342.

CLARKE, H. D., D. SANDERS, M. C. STEWART, AND P. F. WHITELEY (2003): British General Election Study, 2001; Cross-Section Survey. [computer file]. Colchester, Essex: UK Data Archive [distributor], March 2003. SN: 4619.

——— (2006): *British Election Study, 2005: Face-to-Face Survey [computer file]*. Colchester, Essex: UK Data Archive [distributor], November 2006. SN: 5494.

——— (2011): *2010 British Election Study Pre-Election and Post-Election Survey Data*. http://www.bes2009-10.org.

COX, G. W. (1994): "Strategic Voting Equilibria under the Single Nontransferable Vote," *American Political Science Review*, 88(3), 608–621.

——— (1997): *Making Votes Count*. Cambridge University Press, Cambridge.

COX, K. E., AND L. J. SCHOPPA (2002): "Interaction Effects in Mixed-Member Electoral Systems: Theory and Evidence from Germany, Japan, and Italy," *Comparative Political Studies*, 35(9), 1027–1053.

CURTICE, J., AND M. STEED (1988): "Appendix 2: Analysis," in *The British General Election of 1987*, ed. by D. Butler, and D. Kavanagh. Macmillan, Basingstoke.

——— (1992): "Appendix 2: The Results Analysed," in *The British General Election of 1992*, ed. by D. Butler, and D. Kavanagh. Macmillan, Basingstoke.

——— (1997): "Appendix 2: The Results Analysed," in *The British General Election of 1997*, ed. by D. Butler, and D. Kavanagh. Macmillan, Baisingstoke.

DEWAN, T., AND D. P. MYATT (2007): "Leading the Party: Coordination, Direction, and Communication," *American Political Science Review*, 101(4), 827–845.

DIWAKAR, R. (2007): "Duverger's Law and the Size of the Indian Party System," *Party Politics*, 13(5), 539–561.

EVANS, G. (1994): "Tactical Voting and Labour's prospects," in *Labour's Last Chance? The 1992 Election and Beyond*, ed. by A. Heath, R. Jowell, and J. Curtice. Dartmouth, Aldershot.

EVANS, G., AND A. HEATH (1993): "A Tactical Error in the Analysis of Tactical Voting: A Response to Niemi, Whitten and Franklin," *British Journal of Political Science*, 23(1), 131–137.

EVREN, Ö. (2012): "Altruism and Voting: A Large-Turnout Result That Does Not Rely on Civic Duty or Cooperative Behavior," *Journal of Economic Theory*, 147(6), 2124–2157.

FEY, M. (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, AND R. J. JOHNSTON (1996): "Tactical Voting and Party Constituency Campaigning a the 1992 General Election in England," *British Journal of Political Science*, 26, 403–439.

FISHER, S. D. (2004): "Definition and Measurement of Tactical Voting: The Role of Rational Choice," *British Journal of Political Science*, 34(1), 152–166.

——— (2014): "Change in Tactical and Other Motivations for Party Vote Choice," *Working Paper: University of Oxford*.

FISHER, S. D., AND D. P. MYATT (2002a): "Everything is Uncertain and Uncertainty is Everything: Strategic Voting in Simple Plurality Elections," Nuffield College, Oxford: Discussion Paper 115.

——— (2002b): "Tactical Coordination in Plurality Electoral Systems," *Oxford Review of Economic Policy*, 18(4), 504–522.

FRANKLIN, M. N., R. G. NIEMI, AND G. WHITTEN (1994): "The Two Faces of Tactical Voting," *British Journal of Political Science*, 24(4), 549–557.

GAINES, B. J. (1999): "Duverger's Law and the Meaning of Canadian Exceptionalism," *Comparative Political Studies*, 32(7), 835–861.

GOOD, I. J., AND L. S. MAYER (1975): "Estimating the Efficacy of a Vote," *Behavioral Science*, 20(1), 25–33.

GREEN, D. P., AND I. SHAPIRO (1994): *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. Yale University Press, New Haven, CT.

HEATH, A., J. CURTICE, R. JOWELL, G. EVANS, J. FIELD, AND S. WITHERSPOON (1991): *Understanding Political Change: The British Voter 1964–1987*. Pergamon Press, Oxford.

HEATH, A., AND G. EVANS (1994): "Tactical Voting: Concepts, Measurement and Findings," *British Journal of Political Science*, 24(4), 557–561.

HEATH, A., R. JOWELL, AND J. 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, A., R. JOWELL, J. K. CURTICE, J. BRAND, AND J. 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, A., R. JOWELL, J. K. CURTICE, AND P. 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].

HERRMANN, M. (2010): "Voter Uncertainty and Failure of Duvergers Law: An Empirical Analysis," *Public Choice*, 151(1-2), 63–90.

HERRMANN, M., AND F. U. PAPPI (2008): "Strategic Voting in German Constituencies," *Electoral Studies*, 27(2), 228–244.

HOEFFDING, W. (1963): "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association*, 58(301), 13–30.

HOFFMAN, D. T. (1982): "A Model for Strategic Voting," *SIAM Journal of Applied Mathematics*, 42(4), 751–761.

HUGHES, N. (2016): "Voting in Legislative Elections under Plurality Rule," *Journal of Economic Theory*, 166, 51–93.

HUMMEL, P. (2012): "Sequential Voting in Large Elections with Multiple Candidates," *Journal of Public Economics*, 96(3-4), 341–348.

KAWAI, K., AND Y. WATANABE (2013): "Inferring Strategic Voting," *American Economic Review*, 103(2), 624–662.

KSELMAN, D., AND E. NIOU (2010): "Strategic Voting in Plurality Elections," *Political Analysis*, 18(2), 227–244.

LE BRETON, M., D. LEPELLEY, AND H. SMAOUI (2014): "The Probability of Casting a Decisive Vote: From IC to IAC through Ehrhart's Polynomials and m-Dependence," *Working Paper*, (March).

MCKELVEY, R., AND P. ORDESHOOK (1972): "A General Theory of the Calculus of Voting," in *Mathematical Applications in Political Science*, ed. by J. F. Herndon, and J. L. Bernd. University Press of Virginia, Charlottesville, volume 6.

MOSER, R. G., AND E. SCHEINER (2009): "Strategic Voting in Established and Mew Democracies: Ticket Splitting in Mixed-Member Electoral Systems," *Electoral Studies*, 28(1), 51–61.

MYATT, D. P. (2007): "On the Theory of Strategic Voting," *Review of Economic Studies*, 74(1), 255–281.

——— (2015): "A Rational Choice Theory of Voter Turnout," *Working Paper: London Business School*.

——— (2016): "A Theory of Protest Voting," *forthcoming in the Economic Journal*.

MYERSON, R. B., AND R. J. WEBER (1993): "A Theory of Voting Equilibria," *American Political Science Review*, 87(1), 102–114.

NIEMI, R. G., G. WHITTEN, AND M. N. FRANKLIN (1992): "Constituency Characteristics, Individual Characteristics and Tactical Voting in the 1987 British General Election," *British Journal of Political Science*, 22(2), 229–240.

——— (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(4), 549–553.

ORDESHOOK, P. C., AND L. ZENG (1997): "Rational Voters and Strategic Voting: Evidence from the 1968, 1980 and 1992 Elections," *Journal of Theoretical Politics*, 9(2), 167–187.

PALFREY, T. R. (1989): "A Mathematical Proof of Duverger's Law," in *Models of Strategic Choice in Politics*, ed. by P. C. Ordeshook, pp. 69–92. University of Michigan Press, Ann Arbor.

SELB, P. (2012): "Strategic Adaptation to New Electoral Systems," *European Journal of Political Research*, 51(5), 583–606.

SPENKUCH, J. L. (2017): "Expressive vs. Pivotal Voters: An Empirical Assessment," *working paper: Northwestern University*.