

A NEW THEORY OF STRATEGIC VOTING

DAVID P. MYATT / DAVID.MYATT@ECONOMICS.OX.AC.UK

DEPARTMENT OF ECONOMICS, UNIVERSITY OF OXFORD. APRIL 2002.
ADDRESS: ST. CATHERINE'S COLLEGE, OXFORD OX1 3UJ, UNITED KINGDOM

ABSTRACT. Existing game-theoretic analysis of plurality rule elections predicts the complete coordination of strategic voting: A strict interpretation of Duverger's Law. I reach a different conclusion. A group of voters must partially coordinate behind one of two challenging candidates in order to avoid the success of a disliked incumbent. Departing from existing models, the popular support for each challenger is uncertain. Individuals base their votes upon informative signals of candidate support levels. These represent either the social communication of political preferences throughout the electorate, or alternatively the imperfect observation of opinion poll information. The uniquely stable voting equilibrium entails only limited strategic voting and hence partial coordination. This is due to the surprising presence of negative feedback: An *increase* in strategic voting by others actually *reduces* the incentives for an individual to vote strategically. Hence stable multi-candidate support is perfectly consistent with instrumental rationality and fulfilled expectations.

1. DUVERGER'S LAW AND STRATEGIC VOTING

Duverger (1954) introduced his Law by noting that the "simple-majority single-ballot system favors the two-party system." He envisaged an ongoing process involving both voters and political parties with bipartism as an eventual conclusion. Whereas this vision involved only a tendency toward bipartism under the plurality rule, the formal analyses of more recent contributors have generated a rather stricter version of Duverger's Law. Cox (1994) and Myerson and Weber (1993) reinforce Palfrey's (1989) claim that:

"[W]ith instrumentally rational voters and fulfilled expectations, multicandidate contests under the plurality rule should result in only two candidates getting any votes."

This paper is based on Myatt (1999). Elements have appeared previously under the title "Strategic Voting under the Qualified Majority Rule" (Myatt 2000). Stephen D. Fisher inspired this work with his extensive empirical research on tactical voting in Britain, and with many hours of conversation on the topic. I thank colleagues, seminar participants and reviewers for the many helpful comments that have (hopefully) improved the quality of this work since its inception. I remain responsible for any remaining errors.

These authors considered plurality elections where each individual casts a single vote and the candidate with the largest number of votes wins.¹ They found that the uniquely stable equilibrium outcome involves positive support for only two candidates.² This is the result of strategic voting, where an individual may switch her vote away from her preferred candidate. Their Duvergerian prediction is that voters will fully coordinate their strategic behavior. Unfortunately, this strictly bipartite prediction is not borne out by the data. Both the United Kingdom and India provide examples of plurality voting systems with multi-candidate support at a constituency level.³ This might suggest a lack of instrumental rationality on the part of voters. Alternatively, it may point to weaknesses in the specifications of existing formal theories that drive their strictly Duvergerian conclusions.

I will argue that partial coordination of strategic voting is perfectly consistent with stable equilibrium behavior on the part of instrumentally rational individuals. In other words, an appropriately specified formal model predicts only a *tendency* toward bipartism. This argument stems from the observation that existing formal theories assume the perfect common knowledge of the constituency-wide support for different candidates. My response is a model in which voters are uncertain of the preferences of others. They base their decisions on noisy signals of the distribution of preferences throughout the electorate. The analysis shows that strategic voting is a *self-attenuating* rather than *self-reinforcing* phenomenon: An *increase* in strategic voting by others *reduces* the incentives for an individual to vote strategically. This negative feedback pulls strategies away from full coordination and toward multi-candidate support as a stable voting equilibrium.

¹Myerson and Weber (1993) succeed in characterizing equilibria for a wider variety of electoral systems.

²Such equilibria are Duvergerian. Cox (1994) highlighted the existence of non-Duvergerian equilibria. Such equilibria involve switching away from a more to less popular candidate, generating a tie for second place, and thus moderating the incentives for strategic voting. Unfortunately, non-Duvergerian equilibria are highly unstable (Fey 1997) and hence fall outside the class of stable equilibrium predictions.

³Analysis of strategic voting in Britain is provided by Johnston and Pattie (1991), Lanoue and Bowler (1992) and Niemi, Whitten, and Franklin (1992) *inter alia*. New research, based on British Election Study data and analyzing standard intuitive predictions, the bimodality hypothesis of Cox (1994) and the predictions of this paper is reported by Myatt and Fisher (2002). Riker (1976) offered an analysis of the Indian case, and hypothesized that reduced strategic voting was due to the presence of clear Condorcet winner, against which strategic voting is futile. This hypothesis is not supported by the formal analysis presented here.

The argument begins with the following scenario: Imagine a parliamentary constituency or voting district in which a group of dissatisfied voters wish to dislodge a disliked incumbent office-holder. To do so, a proportion $\gamma > \frac{1}{2}$ of the dissatisfied voters must vote in favor of one of two challenging candidates. This is a *qualified majority* voting game: The qualified majority γ must successfully coordinate if the incumbent is to be defeated. This scenario captures many elements of modern plurality elections. In the 1997 British General Election the incumbent (and unpopular) Conservative party polled between $\frac{1}{3}$ and $\frac{1}{2}$ of the votes cast for the three major parties in 270 out of 529 English constituencies. In more than half of England, therefore, anti-Conservative voters needed to successfully coordinate behind either the Labour or Liberal Democrat candidate to ensure a Tory defeat.

What determines the behavior of an instrumental voter in this scenario? She may only influence the outcome of the election with a *casting vote*. This pivotal event occurs when the vote share of one of the challengers is just equal to the required qualified majority γ . An extra vote will then tip the balance away from the disliked incumbent. An instrumental voter then balances the relative probability of the two pivotal events (corresponding to a successful “challenge” by one of the preferred candidates) against her relative preference for the two candidates. This insight is clear from earlier decision-theoretic work by McKelvey and Ordeshook (1972) and Hoffman (1982), and was further explored in a game-theoretic context by Palfrey (1989), Myerson and Weber (1993) and Cox (1994).

Such formal theories successfully highlight these important elements of an instrumental voter’s decision calculus. Unfortunately, the game-theoretic treatments share a common feature: Individual preferences and voting decisions are drawn *independently* from a *commonly known* distribution. It is this, and this alone, that leads to strictly Duvergerian predictions. To see why, consider the relative probability of pivotal events. Mathematically, as the electorate grows large, the probability of a pivotal event involving the leading challenger becomes infinitely larger than the probability of such an event involving the trailing challenger. Equivalently, and perhaps more intuitively, *if* a pivotal event occurs, then it *almost always* involves the leading challenger. Any instrumental voter will, therefore, switch her vote to the leader. Game-theoretic reasoning is unnecessary for a

Duvergerian conclusion: As long as the electorate is large and voting decisions are drawn independently from a commonly known distribution, almost all instrumental voters will switch, even before they account for strategic switching by others.

The fact that voting decisions are drawn independently from a commonly known distribution means that any uncertainty in the Cox (1994), Palfrey (1989) and Myerson and Weber (1993) models is only apparent, and not real. The assumed knowledge of the underlying vote generating process, ensures that such individual-level uncertainty has no affect in the aggregate: As the electorate grows large, the Law of Large Numbers begins to bite, allowing an observer to precisely predict the election's outcome. My aim, therefore, is to remove this potentially undesirable feature and consider seriously the informational underpinnings of the Cox-Palfrey framework.⁴ In particular, I wish to allow voters to be uncertain of the levels of support enjoyed by the candidates.⁵

To achieve this objective, I separate voters' preferences into *common* and *idiosyncratic* components. The common component (reflecting the preferences of the median voter) is shared by all individuals, whereas the idiosyncratic component is distributed independently throughout the electorate. Crucially — and in contrast to Cox-Palfrey — there is *constituency uncertainty* in that the common component is unknown to voters. As the electorate grows large, the idiosyncratic components average out, but constituency uncertainty over the common component remains. This means that *only* uncertainty over the common component matters when considering the realized aggregate preferences (and hence decisions) of a large electorate. Put simply, it is only uncertainty over the identity of the median voter that matters, and yet the Cox-Palfrey modelling paradigm permits only uncertainty over deviations from such a median.

⁴The accurate knowledge required by voters in the Cox-Palfrey framework is fully acknowledged by Cox (1997, p. 78): "A fourth condition necessary to generate pure local bipartism is that the identity of trailing and front-running candidates is common knowledge." In this paper I remove this common knowledge.

⁵The information and beliefs of voters are both concerns of the empirical literature. Heath and Evans (1994) criticize the Niemi *et al* (1992) measure of strategic voting by observing that it does not allow for the possibility that voters are mistaken in their perceptions of the likely chances of the various parties winning the constituency.

An immediate consequence of this approach is that strategic incentives are *finite*. Conditional on the occurrence of a pivotal event, a voter can no longer be sure that the tie involves the leading challenger — simply because she cannot be entirely certain that her perception of the leader’s identity is actually correct. A straightforward consequence is that the presence of constituency uncertainty ensures that the incentives to vote strategically are bounded in large electorates, so that in a decision-theoretic context (when voters do not account for strategic switching by others) a fully Duvergerian outcome is avoided.

This observation by itself does not invalidate the Cox-Palfrey predictions. Their approach was explicitly game-theoretic, where voters take into account strategic switching by others. The opportunity for strict bipartism is created by the familiar logic of *positive feedback*: Strategic voting reduces the support for the less popular challenger. This loss of support (and gain for the leader) enhances the incentive to vote strategically, further eroding the support of the second challenger. This tale of positive feedback leads to the “bandwagon effect” of Simon (1954): *Self-reinforcing* strategic voting expands until the leading challenger attracts all anti-incumbent votes, and a fully Duvergerian outcome is reached as a stable equilibrium. This logic is flawed. When voters do not have a common understanding of the constituency situation, they may be concerned that the bandwagon is rolling in the opposite direction, and will be cautious about switching their vote. Anticipation of more switching by others (a speedier bandwagon) enhances this caution. This tale of *negative feedback* suggests that strategic voting may well be a *self-attenuating* phenomenon.

Investigating this issue, I provide an explicit model of the information sources on which voting decisions are based. Each voter observes a signal of the common component to constituency-wide preferences. Her decision is then based on this signal as well as her preferences. Importantly, her signal helps to determine her beliefs about the behavior of the remaining voters and hence her incentive to vote strategically. Perhaps surprisingly, if all other voters *increase* the response of their behavior to their signal (equivalently, the extent of strategic switching is increased) then the best response of an individual is to *reduce* her response to her own signal in turn. This confirms the negative feedback hypothesis.

Why is this? Consider a benchmark decision-theoretic scenario in which an instrumental individual expects all others to vote truthfully. A relatively large lead in support for candidate 1 may be required to achieve the qualified majority γ , and similarly a relatively large lead for candidate 2 for it to do the same. The instrumental voter compares the probabilities of these events which are *relatively far apart* and will have *different* probabilities of occurrence. The incentive to vote strategically may then be large. Suppose instead that the instrumental voter anticipates that others are likely to vote strategically by responding strongly to their signals. A relatively small lead in the true support for candidate 1 is all that is required to hit the qualified majority of γ . Such a small lead in true support will lead to signals indicating candidate 1's status as leader, and hence strategic switching away from candidate 2. This expands candidate 1's lead, enabling it to reach γ . Identical logic shows that a relatively small lead in true support for candidate 2 is required for it to do the same. The instrumental voter must contemplate two situations involving *relatively smaller* leads. Such events are *relatively close* and hence will have *similar* probabilities of occurrence. Thus the incentive to vote strategically is small.

The self-attenuation of strategic voting may be counter-intuitive, but nevertheless its implications concur with both informal reasoning and observation upon further reflection. A first implication is that appropriate game-theoretic considerations (or increased sophistication on the part of voters) actually *reduce* the impact of strategic voting. When voters consider the strategic decisions made by others, they must also consider the information upon which such decisions are based. This results in the caution that draws voters away from the intensity of strategic switching generated from a decision-theoretic specification. A second implication helps to reconcile formal theory with the observed exceptions to bipartism. Negative feedback leads away from fully coordinated voting behavior, where almost all individuals back the perceived leading challenger. In the qualified majority voting game studied here, the uniquely stable equilibrium outcome entails positive support for both candidates — it exhibits only a *tendency* toward bipartism in the spirit of Duverger's (1954) original legislation. The paper represents, therefore, not so much a "new theory" but rather a better formalization of Duverger's psychological effect.

The argument sketched here is formalized and expanded in the remainder of the paper. I describe the qualified majority voting game, preferences and information sources in Section 2. Section 3 demonstrates that only constituency uncertainty matters in large electorates. I explain the negative feedback phenomenon and characterize the uniquely stable equilibrium in Section 4. The behavior of this equilibrium in response to the constituency situation and the voters' information sources is assessed in Section 5.

2. A MODEL OF QUALIFIED MAJORITY VOTING

2.1. Voting Rules.

An electorate of $n + 1$ anti-incumbent voters is indexed by $i \in \{0, 1, \dots, n\}$. The collective decision $j \in \{0, 1, 2\}$ is taken by *qualified majority voting*. This works as follows. Each individual casts a single vote for either of two challenging candidates $j \in \{1, 2\}$. The disliked incumbent is $j = 0$. Denoting the vote totals for $j \in \{1, 2\}$ as x_1 and x_2 respectively it follows that $x_1 + x_2 = n + 1$. Based on these votes, the winning candidate is:⁶

$$j = \begin{cases} 0 & \max\{x_1, x_2\} \leq \gamma_n n \\ 1 & x_1 > \gamma_n n \\ 2 & x_2 > \gamma_n n \end{cases} \quad \text{where} \quad \gamma_n = \frac{\lceil \gamma n \rceil}{n} \quad \text{and} \quad \frac{1}{2} < \gamma < 1$$

$\gamma > 1/2$ ensures that first, it is impossible for both $j \in \{1, 2\}$ to meet the winning criterion of $x_j > \gamma_n n$, and second, a challenger must have a strict majority of the $n + 1$ strong electorate in order to win. The parameter γ gives a measure of the degree of coordination required to defeat the incumbent. For $\gamma \downarrow \frac{1}{2}$, only a simple majority is required, whereas for $\gamma \uparrow 1$ complete coordination is needed to defeat the incumbent.⁷

The 1970 New York senatorial election, highlighted by Riker (1982), provides a classic example of the scenario described here. The two candidates $j \in \{1, 2\}$ correspond to the liberals Richard L. Ottinger and Charles E. Goodell, whereas the disliked $j = 0$ corresponds

⁶The notation $\lceil y \rceil$ indicates the least integer than it is weakly greater than y .

⁷The interpretation of qualified majority voting as an anti-incumbent coordination problem neglects any strategic behavior by incumbent supporters, and hence helps to focus the model on two-way strategic switching. A similar approach is used elsewhere — see, for instance, Fey (1997).

to the conservative James R. Buckley.⁸ In fact, this latter example gives some guidance for an appropriate choice of γ . Buckley polled 2,288,190 votes (39%) against the combined total of $n + 1 = 3,605,704$ votes (61%) for the opposing candidates. The qualified majority of liberal votes needed to defeat Buckley was approximately $\gamma = 39\%/91\% \approx 63.5\%$.

2.2. Preferences.

Voters are instrumentally rational in the sense that they pursue payoffs that are contingent only on the winning candidate. Voter i receives a payoff u_{ij} when j wins the election. All $n + 1$ voters strictly prefer both $j \in \{1, 2\}$ to the disliked incumbent, yielding payoff normalizations of $u_{i0} = 0$ and $\min\{u_{i1}, u_{i2}\} > 0$. The relative preference for the two challengers varies throughout the electorate. Later analysis will confirm that the ratio of the two payoffs u_{i1} and u_{i2} is sufficient to describe an individual's preferences. Taking logarithms, I define $\tilde{u}_i \equiv \log[u_{i1}/u_{i2}]$ to summarize this. The sign of \tilde{u}_i determines identity of the first preference candidate, and the size $|\tilde{u}_i|$ determines the intensity of this first preference. I break \tilde{u}_i down into two components.

Assumption 1. \tilde{u}_i is decomposed into a common component η and idiosyncratic component ε_i :

$$\tilde{u}_i \equiv \log \left[\frac{u_{i1}}{u_{i2}} \right] = \eta + \varepsilon_i$$

where η is common to everyone and $\varepsilon_i \sim N(0, \xi^2)$, independently throughout the electorate.

A first interpretation is that η represents constituency-wide factors affecting all voters, whereas ε_i represents the idiosyncratic preference of an individual. A second interpretation is that η is the relative preference, and hence essential identity, of the median voter, since $\Pr[\tilde{u}_i \geq \eta] = \Pr[\varepsilon_i \geq 0] = 1/2$. Furthermore, η is also the average voter, since $\eta = E[\tilde{u}_i]$ is the expected log relative preference across voters, conditional on any constituency level information. Finally, η may be inferred from the fraction π of the electorate

⁸This traditional example of a three horse race is used effectively in the recent undergraduate text of Morton (2001). Goodell was an incumbent Republican who had taken a liberal stance on the Vietnam War, and hence received the nomination of the Liberal Party. The New York Conservative Party, however, rather than following the "fusion" route supported Buckley.

with a first preference for candidate 1: $\pi = \Pr[\tilde{u}_i \geq 0] = \Pr[\varepsilon_i \geq -\eta] = \Phi(\eta/\xi)$ where Φ is the cumulative distribution function of the standard normal distribution. This may, of course, be inverted to yield $\eta = \xi\Phi^{-1}(\pi)$. To specify the constituency situation, therefore, I may either specify the identity of median voter η directly or alternatively specify the fraction of voters π who rank candidate 1 first.

The fully parametric specification for the distribution of ε_i , while not critical, permits a simple underlying foundation for the signal specification developed below. It also offers convenience of interpretation. For instance, the variance term ξ^2 provides a measure of the degree of idiosyncrasy throughout the electorate.

2.3. Information.

Whereas voter i is assumed to know her own relative preference \tilde{u}_i , I do *not* allow her to observe the decomposition into common and idiosyncratic components. This means that the median voter η (and hence the proportion $\pi = \Phi(\eta/\xi)$ who favor candidate 1) is unknown to any individual. Of course, individuals will have at least some beliefs about η . I turn, therefore, to consider the information sources on which such beliefs are based.⁹

Voters begin with a common and diffuse prior over η . Equivalently, prior to the receipt of any informative signals they are ignorant of the electoral situation.¹⁰ Tighter beliefs over η are generated following the acquisition of information pertaining to the constituency situation. This is encapsulated in an informative signal δ_i of the common component η .

Assumption 2. *Voter i privately observes an informative signal $\delta_i \sim N(\eta, \kappa^2)$. Conditional on η , this is independently distributed, but may be correlated with the idiosyncratic component ε_i .*

⁹The qualified majority voting game is thus a global game (Carlsson and van Damme 1993) in the sense that it is a game “of incomplete information whose type space is determined by the players each observing a noisy signal of the underlying state” (Morris and Shin 2001). The state variable in this case is η , and the noisy signal of the underlying state is given in Assumption 2 below.

¹⁰An alternative would be to allow voters to begin with a prior belief $\eta \sim N(\mu, \sigma^2)$ and allow $\sigma^2 \rightarrow \infty$. All the results continue to hold with a non-diffuse prior so long as σ is sufficiently large. Adopting a diffuse prior belief does not eliminate the possibility that voters have prior information stemming from earlier elections, or from current media or opinion poll analysis. So long as such information is transmitted with some (perhaps small) noise then it may be viewed within the context Assumption 2.

Following observation of δ_i , voter i updates her diffuse prior to form a posterior belief $\eta \sim N(\delta_i, \kappa^2)$.¹¹ By inspection, it is clear that δ_i represents voter i 's perception of the median voter's identity and $1/\kappa^2$ is the accuracy of this perception. Importantly, different voters receive different signals δ_i , and hence hold different expectations of candidate support. The variance κ^2 also measures the variation in opinions across the electorate. For large κ^2 , we would expect voters to have differing opinions of the constituency situation.

A possible interpretation of Assumption 2 is the social communication of preferences throughout the electorate.¹² Suppose that voter i observes the preferences of $m - 1$ randomly chosen members of the electorate, indexed by k . She also observes her own preference \tilde{u}_i . This sample (of total size m) of preferences provides her with an information source with which to estimate η (or equivalently π).¹³ Given the normality assumption, the sample mean is a sufficient statistic for the sample, generating an aggregate signal δ_i :¹⁴

$$\delta_i = \frac{1}{m} \sum_{k=1}^m \tilde{u}_k \sim N\left(\eta, \frac{\xi^2}{m}\right)$$

Hence an informative signal with variance $\kappa^2 = \xi^2/m$ may be thought of as a detailed "private opinion poll" of size m . The choice of m would correspond to number of people with whom an individual interacts, so long as such people are drawn at random from

¹¹Alternatively, begin with the prior $\eta \sim N(\mu, \sigma^2)$. Bayesian updating (DeGroot 1970) yields:

$$\eta | \delta_i \sim N\left(\frac{\kappa^2 \mu + \sigma^2 \delta_i}{\kappa^2 + \sigma^2}, \frac{\kappa^2 \sigma^2}{\kappa^2 + \sigma^2}\right)$$

Allowing $\sigma^2 \rightarrow \infty$ yields a posterior belief of $\eta \sim N(\delta_i, \kappa^2)$.

¹²Pattie and Johnston (1999) demonstrated that the contextual effects of conversations with family, acquaintances and others were associated with vote-switching behavior in the British General Election of 1992.

¹³I am implicitly assuming that she can elicit the *true* preferences of those within her sample rather than their *stated* preferences. Sampled individuals may, of course, choose to misrepresent their preferences in order to strategically manipulate the beliefs of the recipient. Of course, the recipient would anticipate such manipulation and adjust accordingly. I side-step this issue by supposing that information acquisition occurs over a period of time prior to an election, when individuals in the community have little opportunity or ability to hide their true political preferences.

¹⁴The inclusion of a voter's own preferences within the signal results in correlation between δ_i and ε_i :

$$\delta_i = \eta + \frac{1}{m} \left(\varepsilon_i + \sum_{k \neq i} \varepsilon_k \right) \Rightarrow \text{cov}[\delta_i, \varepsilon_i] = \frac{\text{E}[\varepsilon_i^2]}{m} = \frac{\xi^2}{m} = \kappa^2$$

which yields a correlation coefficient of $\rho = \kappa/\xi$ between ε_i and δ_i . More generally, I allow $\rho \geq \kappa/\xi$ to allow for the possibility that a voter communicates with others who have idiosyncratic components that are correlated with her own.

the population. If voters sample individuals who are similar to them, however, then the effective precision of their information (measured by m) will be dramatically lower.¹⁵

But what about other information sources? An example might well be the widespread publication of opinion polls during an election. In many election scenarios, however, these tend to occur at the *national* level, whereas candidates are elected at a *regional* level. At a regional (i.e. constituency) level opinion polls are rather less common.¹⁶ Nevertheless, the possibility of opinion polls or other central information sources must be taken seriously, and may still be viewed within the context of Assumption 2. For instance, if an opinion poll were to perfectly identify η , then κ^2 would correspond to any (potentially small) noise in a voter's observation of such an opinion poll. In fact, a signal δ_i correctly identifies the leading challenger with probability $\alpha = \Phi(\eta/\kappa)$, and hence α is the accuracy of a voter's observation of the media.¹⁷ Allowing $\kappa^2 \rightarrow 0$ (or equivalently $\alpha \rightarrow 0$) generates the perfect observation of a perfectly revealing public information source.

3. OPTIMAL VOTING BEHAVIOR IN LARGE ELECTORATES

3.1. Optimal Voting Behavior.

Consider the decision of voter $i = 0$. She may only influence the election if she is *pivotal*. This happens if, absent her vote, there is a tie between a candidate and the required qualified majority. To describe pivotal events, I use x to denote the total number of votes cast for candidate 1 among the n other voters $i \geq 1$. If $x = \gamma_n n$, then one more vote will allow candidate 1 to win. Similarly, if $n - x = \gamma_n \Leftrightarrow x = (1 - \gamma_n)n$ then candidate 2 is in the

¹⁵Suppose, for instance, that individuals live within communities, and interact only with members of their own community. Any common cross-community idiosyncratic shock will then generate a tight lower bound to κ^2 , since the sampling procedure cannot eliminate community effects. For instance, if 20% of cross-electorate preference variation is due to variation across communities, then it can be shown that $\kappa^2 \geq \xi^2/5$, or equivalently $m \leq 5$. See Myatt (2002) for more details.

¹⁶Once again, the 1997 UK General Election provides an example. Evans, Curtice and Norris (1998) note that 47 nationwide opinion polls were conducted during the election campaign. By contrast, only 29 polls were conducted in 26 different constituencies at a constituency level, out of a total of 659 constituencies.

¹⁷Since $\kappa = \xi\Phi^{-1}(\pi)$, and $\kappa^2 = \xi^2/m$, then $\alpha = \Phi(\sqrt{m}\Phi^{-1}(\pi))$ or equivalently $m = [\Phi^{-1}(\alpha)/\Phi^{-1}(\pi)]^2$, and hence α may be used as a primitive of the model. Suppose, for instance, that $\pi = 0.6$ and that voters can identify the leading challenger with accuracy $\alpha = 0.85$. This corresponds to a value of $m = 16.8$, and hence is equivalent to the a private sample of approximately 17 individuals.

same position. In both of these settings, $i = 0$ has a casting vote. Conditioning on any information available to her, she will consider the probabilities of the two pivotal events:

$$q_1 = \Pr[x = \gamma_n n] \quad \text{and} \quad q_2 = \Pr[x = (1 - \gamma_n)n]$$

Voting for candidate 1 will generate a positive payoff of u_1 with probability q_1 , and hence generate an expected payoff gain of $q_1 u_1$. Similarly, a vote for candidate 2 yields an expected payoff gain of $q_2 u_2$.¹⁸ These observations lead to the following simple lemma.

Lemma 1. *For voter $i = 0$ with payoffs u_1 and u_2 , an optimal voting rule must satisfy:*

$$\text{Vote} \begin{cases} 1 & q_1 u_1 > q_2 u_2 \\ 2 & q_2 u_2 > q_1 u_1 \\ 1 \text{ or } 2 & q_1 u_1 = q_2 u_2 \end{cases}$$

Notice that whenever a candidate enjoys strong support (when $x > \gamma_n n$ or $1 - x > \gamma_n n$) then a single vote has no effect. It follows that voter $i = 0$ has no interest in such events. For instance, a belief that candidate 1 is very likely to enjoy the vast amount of challenger support does not necessarily attract her vote. She is only interested in outcomes in which a candidate *just* reaches the required qualified majority. When this is possible for both candidates (so that $\min\{q_1, q_2\} > 0$) the following definition may be employed.

Definition 1. *Define the pivotal log likelihood ratio or strategic incentive as $\lambda \equiv \log[q_1/q_2]$.*

Using this definition together with Lemma 1, an instrumental voter should vote for $j = 1$ whenever $u_1 q_1 / u_2 q_2 > 0 \Leftrightarrow \log[u_1/u_2] + \log[q_1/q_2] > 0 \Leftrightarrow \tilde{u} + \lambda > 0$. A voter balances her *relative preference* for the candidates (represented by \tilde{u}) against the *relative likelihood* of her vote influencing the outcome (represented by its logarithm, the strategic incentive λ). When $\lambda = 0$ (so that $q_1 = q_2$) the voter finds each pivotal event to be equally likely, hence faces no strategic incentive and votes straightforwardly for her most preferred candidate.

Lemma 1 describes optimal behavior contingent on her beliefs about pivotal events. A voter will use her signal and her expectation of the strategies used by others to form these

¹⁸When considering the behavior of voter $i = 0$ I omit the subscript i for notational simplicity.

beliefs. At this point I restrict attention to strategies that are contingent only on payoff-relevant information. Such strategies may be categorized as follows.

Definition 2. A symmetric voting strategy $v(\delta_i, \tilde{u}_i) : \mathbb{R}^2 \mapsto [0, 1]$ is the probability of a vote for candidate 1, contingent on the signal and preferences. It is monotonic if it is (weakly) increasing in its arguments, involves strict multi-candidate support if it takes both of the values 0 and 1 for appropriate δ_i and \tilde{u}_i , and is fully coordinated if $v(\delta_i, \tilde{u}_i) = 0 \forall \delta_i, \tilde{u}_i$ or $v(\delta_i, \tilde{u}_i) = 1 \forall \delta_i, \tilde{u}_i$.

Whereas the signal and payoffs are indexed by i , the function $v(\delta_i, \tilde{u}_i)$ is not. It follows that, under symmetric voting strategies, a vote choice is not contingent on an individual's identity, but only on the information available to them. A monotonic voting strategy means that an increase in the preference and signal for an option cannot reduce the probability of a vote for that option. A voting strategy exhibits multi-candidate support if a voter supports each option with certainty, for an appropriate choice of signal and preferences. Finally, a fully coordinated voting strategy involves a definite vote for one of the options independent of the signal and preference realization.

3.2. No Constituency Uncertainty.

What would happen if the common component η (and hence the true support π for candidate 1) were known? Restricting to a symmetric strategy profile among voters $i \geq 1$ (Definition 2), voting decisions are contingent solely on realized payoffs and signals. Conditional on η , these payoffs and signals are independently distributed. It follows that voting decisions are independently distributed. Indeed, I may write $p = [v(\delta_i, \tilde{u}_i) | \eta]$ for the (statistically independent) probability that a randomly selected individual votes for candidate 1.¹⁹ The vote total x for candidate 1 among the n individuals $i \geq 1$ follows a binomial distribution with parameters p and n , yielding pivotal probabilities:

$$q_1 = \Pr[x = \gamma_n n] = \binom{n}{\gamma_n n} p^{\gamma_n n} (1-p)^{(1-\gamma_n)n} \longrightarrow 0 \quad \text{as } n \rightarrow \infty$$

¹⁹Notice that p is the probability of a vote for candidate 1, whereas $\pi = \Pr[\tilde{u} \geq 0 | \eta]$ is the probability that a voter's first preference is for candidate 1. In the absence of any strategic voting $p = \pi$.

and similarly for q_2 . Hence, for large constituencies, the *absolute* probability of a pivotal outcome falls to zero. As I have argued, it is not the absolute but rather the *relative* likelihood of pivotal events, captured by the strategic incentive λ , that determine voting behavior. Taking the position of the voter $i = 0$, I may evaluate λ to obtain:

$$\lambda = \log \left[\frac{q_1}{q_2} \right] = \log \left[\frac{p^{\gamma n} (1-p)^{1-\gamma n}}{p^{1-\gamma n} (1-p)^{\gamma n}} \right]^n \longrightarrow \begin{cases} +\infty & p > 1/2 \\ -\infty & p < 1/2 \end{cases} \quad \text{as } n \rightarrow \infty \quad (1)$$

Unless $p = 1/2$, the strategic incentive λ grows without bound as the electorate grows large. If voter $i = 0$ acts optimally, she will almost always vote for candidate 1 whenever $p > \frac{1}{2}$. Extending this optimal response to other voters, there is complete coordination, with the entire electorate strategically abandoning option 2. Notice that, in this setting, the strategic incentive λ is *entirely* driven by idiosyncratic uncertainty, and that game-theoretic reasoning is unnecessary for a strictly Duvergerian outcome.

3.3. Uncertain Common Effect.

But what if η (the identity of the median voter) is uncertain? Conditional on η , voting decisions continue to be binomial with parameters p and n . But since p depends on η , and η is unknown, it follows that p is unknown. I focus on monotonic voting strategies (Definition 2). If the voting strategy profile adopted by others is fully coordinated, then irrespective of η there is no possibility of a pivotal outcome.²⁰ Next, consider monotonic voting strategies that exhibit multi-candidate support. It is immediate that $p = E[v(\delta_i, \tilde{u}_i | \eta)]$ is a continuous function of η . Given the signal specification, δ is a sufficient statistic for η . I write $f(p | \delta)$ for the density of her beliefs over p , conditional on δ .

Lemma 2. *Suppose that $v(\delta_i, \tilde{u}_i)$ is monotonic and exhibits strict multi-candidate support. Defining $p = E[v(\delta_i, \tilde{u}_i | \eta)]$, the density $f(p | \delta)$ is continuous and strictly positive on $(0, 1)$.*

Proof. See Appendix A. □

²⁰This, of course, means that the instrumental incentives for a voter will be exactly zero, rather than zero in the limit as $n \rightarrow \infty$. In other words, if an instrumental voter expects all others to fully coordinate with probability one, then instrumental considerations will be no guide to her vote choice.

With this in hand, the pivotal probability of a tie involving candidate 1 becomes:

$$q_1 = \int_0^1 \binom{n}{\gamma_n n} [p^{\gamma_n} (1-p)^{1-\gamma_n}]^n f(p|\delta) dp \quad (2)$$

and similarly for q_2 . The binomial probability term in the integrand represents *idiosyncratic* uncertainty. Even with full knowledge of p , the decision of any particular voter is unknown. The second term $f(p|\delta)$ represents *constituency* uncertainty. From the perspective of voter $i = 0$, the constituency-wide support for candidate 1 (represented by p) is unknown. Now, as $n \rightarrow \infty$, it is clear that the probability q_1 vanishes to zero. Its asymptotic properties are interesting, however, and are recorded in the following proposition.

Proposition 1. *Under the conditions of Lemma 2, the pivotal probabilities satisfy:*

$$\lim_{n \rightarrow \infty} (n+1)q_1 = f(\gamma|\delta), \quad \lim_{n \rightarrow \infty} (n+1)q_2 = f(1-\gamma|\delta), \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{q_1}{q_2} = \frac{f(\gamma|\delta)}{f(1-\gamma|\delta)}$$

Sketch Proof. Examine Equation 2 and notice as n grows (and $\gamma_n \rightarrow \gamma$) the integrand becomes increasingly peaked around a maximum at $p = \gamma$. Equivalently, in a large electorate, candidate 1 can only match the required qualified majority when $p = \gamma$. It follows that only density local to γ is relevant in the integral, and hence I may replace $f(p|\delta)$ with $f(\gamma|\delta)$. Bringing this outside the integral, the remaining expression (a Beta density) integrates to $1/(n+1)$, yielding Proposition 1. Appendix A has a formal proof. \square

This proves that in a large constituency, the relative likelihood of the two pivotal events is the relative likelihood that their respective constituency-wide support levels (represented by p and $1-p$) coincide with the critical value γ . Importantly, then, it is only uncertainty over p (generated by uncertainty over the common effect η) that matters. Why is this? As the constituency grows large, the individual idiosyncrasy in payoff and signal realizations are averaged out. Uncertainty over the common component η , however, *cannot* be averaged out and hence becomes the key determinant. In the earlier Cox-Palfrey models only idiosyncratic uncertainty was present. But in the presence of constituency uncertainty, its effect disappears. The results of earlier work, therefore, may well be driven by the wrong source of uncertainty. An immediate implication is that the limiting strategic

incentive λ is *finite*, and hence voter $i = 0$ may well find it optimal to vote for either candidate. Of course, this leaves upon the possibility that strategic voting may be self-reinforcing, leading to an *equilibrium* outcome that involves full coordination — an issue that I address in Section 4.

3.4. Qualified Majority Voting in Large Electorates.

My next step is analyze the qualified majority voting game in a large electorate ($n \rightarrow \infty$). Of course, if a continuum of voters is specified directly, then a single vote can have no effect. Equivalently, the probability of a pivotal outcome and hence the payoff to any vote vanish to zero as the electorate grows large. I have argued, however, that it is not the *absolute* values of payoffs and probabilities that are of importance, but rather their *relative* size. Indeed, in the finite population qualified majority voting game, I may re-scale the payoffs by multiplying through by the electorate size $n + 1$. The payoff for a vote for candidate j becomes $(n + 1)q_j u_j$. Taking the limit as $n \rightarrow \infty$, the payoffs for the limit game (a continuum of voters) are then well defined. There are two cases to consider.

For the first case, suppose that voters $i \neq 0$ adopt a fully coordinated symmetric strategy profile (Definition 2). I define the payoffs for the focal voter $i = 0$ from votes for options 1 and 2 as $U_1 = U_2 = 0$. For the second case, suppose that voters $i \neq 0$ adopt a symmetric and monotonic voting strategy profile that exhibits multi-candidate support. Note that $(n + 1)q_1 \rightarrow f(\gamma | \delta)$ and $(n + 1)q_2 \rightarrow f(1 - \gamma | \delta)$, by Proposition 1. I thus define the payoffs in the limit game as follows:

$$U(\text{Vote 1} | v(\delta_i, \tilde{u}_i), \delta) \equiv f(\gamma | \delta)u_1$$

$$U(\text{Vote 2} | v(\delta_i, \tilde{u}_i), \delta) \equiv f(1 - \gamma | \delta)u_2$$

It is worth observing that re-scaling the payoffs by the electorate size may in fact reflect an appropriate specification of instrumental payoffs in a large electorate. A familiar critique of instrumentally-motivated voting models is that all votes are wasted in a large electorate. Indeed, Proposition 1 confirms that $q_j \rightarrow 0$ at rate $1/(n + 1)$ as $n \rightarrow \infty$. An expansion of the electorate size, however, may also change the size of the payoffs received

by a voter in the event of a win by each of the challenging candidates. If the electorate size $(n + 1)$ were to measure the importance of the election in the eyes of a voter, then $(n + 1)u_j$ would be an appropriate specification of the payoff when j wins the election. In other words, instrumental voters may care more about more important elections.

4. OPTIMAL VOTING AND EQUILIBRIUM

I now focus exclusively on the qualified majority voting game in an unboundedly large electorate, envisaging continuum of voters with payoffs as specified in Section 3.4.

4.1. Linear Voting Strategies.

I begin with the optimal behavior of the focal voter $i = 0$ when the remaining voters adopt a symmetric profile of monotonic voting strategies exhibiting multi-candidate support.

Lemma 3. *If all voters $i \neq 0$ adopt a symmetric monotonic voting strategy exhibiting strict multi-candidate support (Definition 2), then the best response for voter $i = 0$ is to use a linear voting strategy: She votes for candidate 1 whenever $\tilde{u} + a + b\delta \geq 0$, for some $a \in \mathbb{R}$, $b \in \mathbb{R}_+$.*

Sketch Proof. When the n other members of the electorate employ a voting strategy that is positively related to signals and preferences, then there will be two values $\eta_1 > \eta_2$ of the common component that result in $p = \gamma$ and $p = 1 - \gamma$ respectively. Voter $i = 0$ will, therefore, consider the log likelihood ratio of the events $\eta = \eta_1$ and $\eta = \eta_2$. Of course, posterior beliefs over η are normal with a mean of δ . Log likelihood ratios of the normal are linear in its mean, and hence:

$$\lambda(\delta) = \log \left[\frac{f(\gamma | \delta)}{f(1 - \gamma | \delta)} \right] = a + b\delta$$

for some a and $b > 0$. Voter $i = 0$ then finds it optimal to vote for candidate 1 whenever $\tilde{u} + \lambda(\delta) \geq 0$,²¹ which is equivalent to $\tilde{u} + a + b\delta \geq 0$. See Appendix A for a full proof. \square

This ensures that so long as other voters react positively to both their payoffs and private information, then I may focus on the class of simple linear strategies. The linearity of

²¹Without loss of generality, I assume that she votes for candidate 1 when indifferent.

optimal voting strategies has further implications: First, the class of symmetric and monotonic voting strategies exhibiting multi-candidate support is closed under best response, since the linear strategies of Lemma 3 are within this class. Second, the class of *linear* voting strategies is closed under best response. Third, any symmetric and monotonic voting equilibrium exhibiting multi-candidate support must involve linear strategies.

To say more about the nature of optimal voting strategies, and the existence and nature of equilibrium strategies, requires the exact properties of the parameters a and b .

Lemma 4. *The class of symmetric linear voting strategy profiles is closed under best response. If all voters $i \neq 0$ adopt a linear voting strategy $v(\delta_i, \tilde{u}_i) = 1 \Leftrightarrow \tilde{u}_i + a + b\delta_i \geq 0$, then a best response for the focal voter $i = 0$ is to adopt a linear voting strategy $v(\delta_i, \tilde{u}_i) = 1 \Leftrightarrow \tilde{u} + \hat{a} + \hat{b}\delta \geq 0$, where:*

$$\hat{a}(a, b) = \frac{\hat{b}a}{1+b}, \quad \hat{b}(b) = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)}{\kappa^2(1+b)} \quad \text{where} \quad \tilde{\kappa}^2 = \text{var}[\varepsilon_i + b\delta_i \mid \eta] \quad (3)$$

Proof. See Appendix A. □

Linear strategies are easily interpreted. The parameter a is the strategic incentive faced by a voter following receipt of a neutral signal $\delta = 0$. It is a signal-independent *bias* toward candidate 1 (when $a > 0$) or candidate 2 (when $a < 0$). From Lemma 4 voter $i = 0$ will exhibit a bias (e.g. $\hat{a} > 0$) if and only if she expected the rest of the population to exhibit such a bias ($a > 0$). In contrast, b represents the response of the strategic incentive to a voter's private signal. The shape of $\hat{b}(b)$ shows how voter $i = 0$ use of her signal changes in response to increased ($b \uparrow$) or decreased ($b \downarrow$) responsiveness on the part of others.

4.2. Strategic Voting and Negative Feedback.

I have suggested that strategic voting may exhibit *negative* feedback. Since $a = 0$ in a stable voting equilibrium (see below) the appropriate vehicle for evaluating this hypothesis is b . As observed above, this evaluation is drawn from an inspection of $\hat{b}(b)$.

Lemma 5. *$\hat{b}(b)$ is decreasing from $\hat{b}(0) = 2\xi\Phi^{-1}(\gamma)/\kappa^2$ to $\lim_{b \rightarrow \infty} \hat{b}(b) = 2\Phi^{-1}(\gamma)/\kappa$.*

Proof. See Appendix A. □

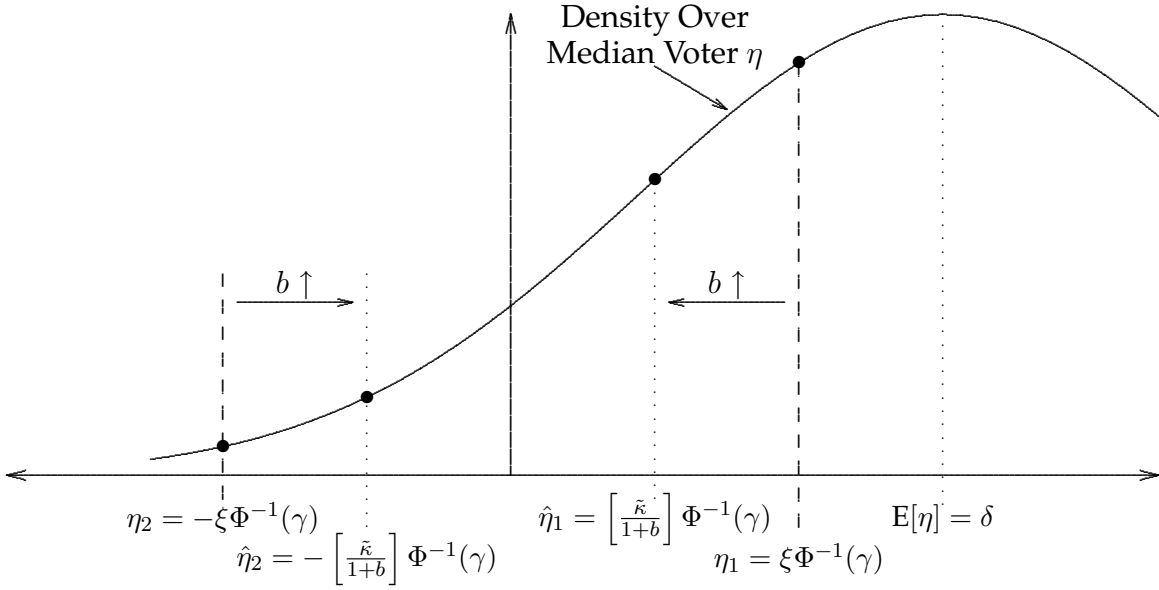


FIGURE 1. Explaining Negative Feedback

Lemma 5 states that $\hat{b}(b)$ is *decreasing* in b : An *increase* in the tendency by others to vote strategically ($b \uparrow$) *reduces* the tendency of voter $i = 0$ to vote strategically ($\hat{b} \downarrow$). Figure 1 helps to explain this phenomenon. Suppose that all voters $i \neq 0$ vote according to their true first preference (equivalent to $a = 0$ and $b = 0$) and that voter $i = 0$ receives a signal δ . Her posterior belief over the identity of the median voter η is distributed around δ , as illustrated by the density in Figure 1. The support for candidate 1 is $p = \Pr[\eta + \varepsilon_i \geq 0] = \Phi(\eta/\xi) \Leftrightarrow \eta = \xi\Phi^{-1}(p)$. This candidate will just reach the required qualified majority γ when $p = \gamma$, or equivalently when $\eta = \eta_1 = \xi\Phi^{-1}(\gamma)$. Similarly, candidate 2 will just reach the qualified majority when $\eta = \eta_2 = -\xi\Phi^{-1}(\gamma)$. Voter $i = 0$ will compare the relative likelihood of the events $\eta = \eta_1$ and $\eta = \eta_2$. It is clear from Figure 1 that $\eta = \eta_1$ is rather more likely, and hence she faces a large incentive to vote strategically for candidate 1.

Suppose instead that all voters $i \neq 0$ are willing to vote strategically by responding strongly to their private signals, so that $b > 0$. The support for candidate 1 is now:

$$p = \Pr[\tilde{u}_i + b\delta_i \geq 0] = \Phi\left(\frac{(1+b)\eta}{\sqrt{\text{var}[\tilde{u}_i + b\delta_i | \eta]}}\right) = \Phi\left(\frac{(1+b)\eta}{\tilde{\kappa}}\right) \Rightarrow \eta = \frac{\tilde{\kappa}\Phi^{-1}(p)}{1+b}$$

where $\tilde{\kappa}$ is taken from Equation 3. Once again, voter $i = 0$ assesses the likelihood of candidate 1 just reaching the qualified majority γ . This happens when $\eta = \hat{\eta}_1 = \tilde{\kappa}\Phi^{-1}(\gamma)/(1+b)$. This is a less extreme criterion for the median voter to reach than previously. Formally:

$$\hat{\eta}_1 = \frac{\tilde{\kappa}\Phi^{-1}(\gamma)}{1+b} < \xi\Phi^{-1}(\gamma) = \eta_1$$

where the inequality follows from the observation that $\tilde{\kappa} < (1+b)\xi$.²² By the same logic, candidate 2 will just reach γ when $\eta = \hat{\eta}_2 = -\tilde{\kappa}\Phi^{-1}(\gamma)/(1+b) > \eta_2$. Voter $i = 0$ will assess the relative likelihood of $\eta = \hat{\eta}_1$ versus $\eta = \hat{\eta}_2$. Clearly (Figure 1) these two values are much closer together. It is no longer the case that candidate 1 is much more likely to contend for the qualified majority, and reducing the incentive to vote strategically.

Although negative feedback may seem counter-intuitive, it accords with intuition upon further reflection. When b is high, individuals respond strongly to their signals, increasing the likelihood of a strategic vote. Importantly, however, it increases the probability of a strategic vote in *both* directions. Voter $i = 0$ with signal $\delta > 0$ is concerned that other voters may observe signals $\delta_i < 0$, yielding a pivotal outcome involving candidate 2 rather than candidate 1. For high δ , this event seems most unlikely — surely candidate 1 will almost certainly win? But if candidate 1 will almost certainly win, then the vote of $i = 0$ has no effect. She can *only* influence the outcome when there is a tie. But if there is a tie, then her strong signal must have overstated the support for candidate 1. She must therefore envisage a much lower true value for η . It is then more reasonable for her to consider true values of the common component (or median voter) satisfying $\eta < 0$.

4.3. Stable Voting Equilibria.

I now seek symmetric and monotonic strategic voting equilibria. The first possibility that I wish to consider is equilibria exhibiting strict multi-candidate support. Following Lemma 3, such equilibria must involve linear voting strategies, where individual i votes for candidate 1 if and only if $\tilde{u}_i + a + b\delta_i \geq 0$. Of course, if all voters use such a strategy

²²This inequality is equivalent to $\tilde{\kappa}^2 < (1+b)^2\xi^2$. By definition $\tilde{\kappa}^2 = \text{var}[\varepsilon_i + b\delta_i | \eta] = \xi^2 + b^2\kappa^2 + 2b\rho\xi\kappa$ where ρ is the correlation coefficient between ε_i and δ_i . The desired inequality holds whenever $2b(\rho\kappa\xi - \xi^2) < b^2(\xi^2 - \kappa^2)$. This holds because the left hand side is negative, since $\rho \leq 1$, and $\kappa^2 \leq \xi^2$ by assumption.

then it is a best response for voter $i = 0$ to use a linear strategy with coefficients $\hat{a}(a, b)$ and $\hat{b}(b)$ from Lemma 4. For an equilibrium I require a pair $\{a^*, b^*\}$ satisfying:

$$a^* = \hat{a}(a^*, b^*) = \frac{b^* a^*}{1 + b^*} \quad \text{and} \quad b^* = \hat{b}(b^*) = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)}{\kappa^2(1 + b^*)}$$

The second equation involves only b and hence may be considered independently.

Lemma 6. *The mapping $\hat{b}(b)$ has a unique fixed point $b^* > 0$. For $\rho \geq \kappa/\xi$, this satisfies:*

$$2\Phi^{-1}(\gamma) \leq \kappa b^* \leq \Phi^{-1}(\gamma) \left\{ 1 + \sqrt{\frac{\Phi^{-1}(\gamma) + 2\xi}{\Phi^{-1}(\gamma)}} \right\}$$

where ρ is the correlation coefficient between δ_i and ε_i . For the social communication interpretation where $\rho = \kappa/\xi$ (see Section 2.3) the bound may be refined to:

$$2\Phi^{-1}(\gamma) \leq \kappa b^* \leq \Phi^{-1}(\gamma) \sqrt{2 + 2\sqrt{\frac{(\Phi^{-1}(\gamma))^2 + \xi^2}{(\Phi^{-1}(\gamma))^2}}} \quad (4)$$

where κb^* attains the upper bound as $\kappa^2 \rightarrow 0$.

Proof. The uniqueness of b^* follows from the fact that $\hat{b}(b)$ is decreasing. The bounds described above are calculated in Appendix A. \square

Turning attention back to a^* , the fixed point equation $a^* = b^* a^* / (1 + b^*)$ has a unique solution at $a^* = 0$. This observation, together with Lemma 6 proves the following proposition.

Proposition 2. *There is a unique symmetric and monotonic voting equilibrium exhibiting multi-candidate support. It involves linear voting strategies and hence only partial strategic voting.*

Proposition 2 demonstrates that partial coordination (and hence multi-candidate support) is perfectly consistent with equilibrium behavior on the part of instrumentally rational voters. Other interesting observations are available. The fact that $a^* = 0$ means that there can be no systematic bias toward one candidate: All strategic voting is driven by the response of voters to their informative signals of the constituency situation, as measured by $b^* > 0$. Of course, this is a response to a voter's *signal* of the constituency situation

rather than the *actual* situation. For $\eta > 0$, the *realization* of the signal for a particular voter i may well satisfy $\delta_i < 0$. This yields a strategic incentive for voter i to switch *away* from the more preferred option. It follows that strategic voting is bi-directional, with some voters switching in the *wrong* direction.²³

Before moving further, a potential equilibrium selection problem must be addressed. Although the equilibrium described above is unique in the class of monotonic equilibria with multi-candidate support, two other equilibria also exist, if only technically. These are *strict* Duvergerian equilibria, where the entire electorate coordinates on a single candidate without reference to their payoff or signal realizations. In this case, the probability of a pivotal outcome is always zero (in fact the limit game specifies payoffs of $U_1 = U_2 = 0$) and hence all instrumental voters will be exactly indifferent between the two candidates. Exploiting this indifference rather heavily, it is an equilibrium for them to all coordinate on a single candidate. Such strict Duvergerian equilibria are, in the context of the present model, highly unsatisfactory and for a number of reasons. I highlight three critiques here. My first critique is that the indifference of voters would allow non-instrumental motivations to dominate. For instance, suppose that voters were to be equipped with the following lexicographic preference structure: They are primarily instrumental, pursuing instrumental motivations *unless* the instrumental incentives are zero, in which case they vote for their truly favorite candidate. This tiny modification to the specification of the model eliminates any strictly Duvergerian equilibrium. It has no effect, however, on the partially coordinated equilibrium described in Proposition 2.²⁴

My second critique is that a strictly Duvergerian equilibrium requires some form of explicit coordination among voters. The source of such a coordination device is unclear. One

²³Anecdotally at least, this phenomenon was observed in the British General Election of 1997. The disliked Conservative incumbent Michael Howard stood for re-election in the Folkestone and Hythe constituency. Strategic voting was reputed to have occurred in *both* directions. Michael Howard retained his seat, polling 39.0% of the vote. The left-wing parties split the anti-Conservative vote almost exactly — Labour 24.9% and Liberal Democrat 26.9%. Thanks are due to Steve Fisher for this information.

²⁴A response to this critique might be that non-instrumental motivations would take over in the partially coordinated equilibrium, since the absolute probability of a pivotal event falls to zero with n . A defence is the argument of Section 3.4: Instrumental payoffs may scale up with the size of the electorate, offsetting the fall in the absolute probability of a pivotal outcome.

possibility is that all voters perfectly observe a coordinating announcement by a single individual or organization. But how could voters be sure that they are all seeing the same announcement? More subtly, even if all voters see the same coordinating announcement, and moreover see that others are doing the same, this does *not* necessarily mean that there is common knowledge of such an announcement.²⁵ In fact, an instrumental voter will always be interested in exactly those circumstances in which the announcement fails, since it is only in such circumstances that her vote will have any effect. This leads consideration back to equilibria in which voting behavior is responsive to signals and preferences.

My third critique is that the partially coordinated equilibrium of Proposition 2 is uniquely stable in an appropriately specified dynamic, whereas strictly Duvergerian equilibria are not. Specifically: Beginning from any symmetric and monotonic voting strategy with strict multi-candidate support, a sequence of updated best responses will always lead back toward the partially coordinated equilibrium. To understand this idea, begin with the initial hypothesis that all individuals vote straightforwardly for their preferred option. This is equivalent to employing a linear voting strategy with $a_0 = b_0 = 0$. Voter $i = 0$, acting optimally in response to this strategy profile, will employ a linear voting strategy with parameters $a_1 = \hat{a}(0, 0)$ and $b_1 = \hat{b}(0)$. Of course, she may well anticipate a similar response in the electorate at large, and hence update once more to obtain another linear strategy $a_2 = \hat{a}(a_1, b_1)$ and $b_2 = \hat{b}(b_1)$. This thought experiment describes an iterative best response process within the class of linear voting strategies.²⁶ Of course, a starting point within the class of symmetric and monotonic voting strategies with strict multi-candidate support will enter this class within one step (Lemma 3). Formally:

²⁵Such common knowledge is an extreme form of public identification: Voters "...desert the (publicly identified) trailing candidates in order to focus on the (publicly identified) front-running candidates." (Cox 1997, p. 78, his parentheses) The recent economic literature on global games (Morris and Shin 2001) demonstrates that the lifting of common knowledge assumptions has a dramatic effect on the nature of equilibria. Cox (1997) compares the present scenario to the classic "Battle of the Sexes" game, with two coordinated pure strategy and a single mixed strategy Nash equilibrium. In the present context, these correspond to Duvergerian and non-Duvergerian equilibrium. As the seminal contribution of Carlsson and van Damme (1993) demonstrates, a slight weakening of the common knowledge assumption in Battle of the Sexes results in a *unique* Bayesian Nash equilibrium. Exactly the same procedure happens here.

²⁶Earlier work by Fey (1997) considered a process of repeated elections, beginning with an election in which agents act truthfully. The iterative best response process here is viewed as a thought experiment prior to the act of voting, and hence multiple elections are unnecessary.

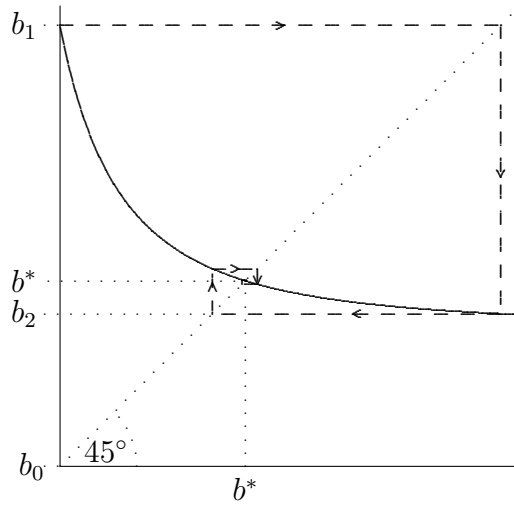


FIGURE 2. $\hat{b}(b)$ and Convergence to Equilibrium

Definition 3. Define the iterative best response process by $b_t = \hat{b}(b_{t-1})$, $a_t = \hat{a}(a_{t-1}, b_{t-1})$.

Having defined this process, global stability may be ascertained.²⁷ The mapping $\hat{b}(b)$ and associated process $\{b_t\}$ are not contingent on a_t , and hence may be considered in isolation.

Lemma 7. b^* is globally stable in the iterative best response dynamic: $b_t \rightarrow b^*$ as $t \rightarrow \infty$.

Proof. See Appendix A. □

Whereas the formal proof is algebraically tedious, a diagrammatic illustration is available. Figure 2 is a plot of \hat{b} , incorporating a path of convergence to the fixed point b^* . The cyclic behavior is a consequence of negative feedback. Begin with $b_0 = 0$, and note that $b_1 = b(0) > 0$. Taking the next step, voter $i = 0$ recognizes the strategic behavior of others. This attenuates her response to her signal, reducing b . Of course, this behavior leaves open the possibility of a limit cycle in the iterative best response process. Lemma 7 ensures that the cycle dampens down, eventually converging to the unique fixed point b^* . The stability of a^* is straightforward. I obtain the following proposition.

²⁷Of course, the iterative best response process cannot begin at a fully coordinated voting strategy profile. Facing such a strategy, an instrumental voter faces no incentives at all! Assuming that such a voter retained the initially postulated strategy profile (in an extension to Definition 3) would allow the fully coordinated Duvergerian equilibria to be (technically) stable in the iterative best response dynamic. Once again, however, this would be crucially dependent upon the exact common knowledge of the initial conditions and the absolute absence of any non-instrumental concerns.

Proposition 3. *From any monotonic symmetric voting strategy with multicandidate support, the iterative best dynamic converges to the globally stable partially coordinated voting equilibrium.*

Proof. Given that $b_t \rightarrow b^*$, it follows that $b_{t+1}/(1 + b_t) \rightarrow b^*/(1 + b^*)$. From this, the global stability of \hat{a}^* is immediate: $a_{t+1}/a_t \rightarrow b^*/(1 + b^*) < 1$ and so $a_t \rightarrow 0$. As the iterative best response process continues, any “bias” toward a particular candidate is eliminated. Combining this observation with Lemma 7 completes the proof. \square

Hence a thought process by which voter $i = 0$ iteratively assesses the likely behavior of others leads to a partially coordinated equilibrium. In particular, this is true when a voter begins with the initial hypothesis that others will act truthfully ($a_0 = b_0 = 0$). No explicit coordinative device is necessary: Iterative reasoning alone allows an instrumental voter to reach a partially (but not fully) coordinated equilibrium voting strategy.

Importantly, the partially coordinated equilibrium described here is *not* related in *any way* to the non-Duvergerian equilibria highlighted by Cox (1994). Understanding the stark differences is important. Cox’s (1994) model requires the *exact* knowledge of candidate support levels, or equivalently the absence of any constituency support. Of course, if the voting strategy generates $p \neq 1/2$ then such a specification would yield an infinite strategic voting incentive and hence the complete coordination of strategic voting (see Equation 1 in Section 3.2). To avoid this, and hence generate multi-candidate support, p needs to be close (and arbitrarily close in a large electorate) to $1/2$. To achieve this, a precisely calculated proportion of the voters who prefer the leading challenger (so for $\pi > 1/2$ voters who view candidate 1 as most preferred) must switch *away* from this leader and toward the trailing candidate. Specifically, if a fraction $\pi - 1/2$ voters do so, then a tie between the challengers ($p = 1/2$) will result in the attenuation of any strategic voting incentives.²⁸ This feature leads to explicit *mis*-coordination: Some voters switch away from a leading challenger in order to stop the strategic incentives that would lead to the successful defeat of the disliked status quo. This is especially perplexing when it

²⁸In fact, for the case $\pi > 1/2$ a slightly greater proportion are required to switch in order to generate a (finite) strategic incentive to switch away from candidate 1.

is understood that such non-Duvergerian equilibria require precise common knowledge of the constituency situation: π or equivalently η must be commonly known by all voters.²⁹ These “knife-edge” properties were recognized by Palfrey (1989), who discounted non-Duvergerian equilibria.³⁰ Cox’s response (1994, p. 609) was that such equilibria “are not unusual in the mathematical sense of being non-generic.” In the present model, the unique partially coordinated equilibrium involves strategic switching toward the leading candidate. It cannot, therefore, exhibit a tie between the challengers. This continues to be true in the limit as $\kappa^2 \rightarrow 0$ and all uncertainty is removed. It follows that non-Duvergerian equilibria *are* non-generic: They only exist when κ^2 is *exactly* equal to zero.

Formally, the partially coordinated stable equilibrium of the present model is radically different from the Cox (1994) non-Duvergerian equilibrium, and yet incorporates *precisely* the informal interpretations sometimes offered for such non-Duvergerian equilibria and yet explicitly excluded from the Cox-Palfrey framework. For instance, Cox (1997, p. 86) interprets the constituency result for the Ross and Cromarty constituency (where the Liberal and Labour candidates almost tied, permitting a Conservative win) from the 1970 UK General Election as a potential non-Duvergerian equilibrium. He claimed that this interpretation requires that “it was not clear who was in third and who in second” and “[N]either [challenger] suffered from strategic desertion.” These are exactly *opposite* to the requirements of non-Duvergerian equilibria which require *exact* knowledge of the gap between parties and *precisely calculated* strategic swing. This intuition, however, is consistent with the partially coordinated equilibrium described here. Voters are indeed uncertain of the ranking of challenging candidates, and this results in lower strategic incentives.

²⁹The exact amount of “wrong way switching” that is required is sensitive to n , and hence outside the context of the limiting voting game ($n \rightarrow \infty$) non-Duvergerian equilibria require common knowledge of the exact electorate size n .

³⁰It is perhaps unsurprising, therefore, that Fey’s (1997) repeated election dynamic diverges away from non-Duvergerian equilibria. Interestingly, his dynamic is rather different from the one described here. In particular, it is path dependent: The limiting Duvergerian equilibrium to which it converges depends on the starting point. Repeated elections are *not* necessary to employ the iterative best response process. This can be simply regarded as a “fictitious play” thought process in the mind of a player.

5. COMPARATIVE STATICS

The first measure of strategic voting that I consider is a voter's response b to her signal of the constituency situation. Both decision-theoretic and game-theoretic cases may be considered via an examination of the best response mapping $\hat{b}(b)$ is the vehicle, where:

$$\hat{b}(b) = \frac{2\Phi^{-1}(\gamma)\sqrt{\text{var}[\varepsilon + b\delta | \eta]}}{\kappa^2(1+b)} = \frac{2\Phi^{-1}(\gamma)\sqrt{\xi^2 + b^2\kappa^2 + 2b\rho\kappa\xi}}{\kappa^2(1+b)} \Rightarrow \hat{b}(0) = \frac{2\Phi^{-1}(\gamma)\xi}{\kappa^2} \quad (5)$$

$\hat{b}(0)$ is the response of a decision-theoretic strategic voter — a voter who acts instrumentally while expecting others to vote for their true first preference. b^* is the response of a game-theoretic strategic voter, who takes account of strategic voting by others. Of course, if $\hat{b}(b)$ is increasing in an exogenous parameter for all b , then (since \hat{b} is decreasing in b) so must the fixed point b^* as well as $\hat{b}(0)$.

Examining Equation 5 it is immediate that \hat{b} is increasing in γ : A voter responds more strongly to her signal when greater coordination is needed. There is a sense in which γ represents the “safety” of a disliked incumbent office-holder. For instance, when $\gamma \rightarrow 1$ full coordination is needed to avoid the status quo. This may run against initial intuition: For English parliamentary constituencies, it means that relatively safe Tory seats will encourage anti-Conservative voters to coordinate their actions. Of course, an instrumental voter cares only about influencing the outcome, and realises that for large γ coordination behind the leading candidate is crucial in such a scenario. Inspection of Equation 5 also reveals that \hat{b} is decreasing in κ^2 . These observations generate the following lemma.

Lemma 8. *A voter's willingness to vote strategically b by responding to her private signal δ is increasing in the qualified majority required γ and the precision of the informative signal $1/\kappa^2$.*

Proof. Inspection of Equation 5. □

Equation 5 also shows that \hat{b} is decreasing in the idiosyncrasy of the electorate ξ . This comparative static may be misleading, however. Fixing the proportion of the population $\pi = \Phi(\eta/\xi)$, the median voter η satisfies $\eta = \xi\Phi^{-1}(\pi)$. But this means that, for fixed π , $|\eta|$ and hence the size of the signal δ that a voter is likely to receive are increasing in ξ . When

the electorate is relatively idiosyncratic (high ξ^2) then voters responds more sluggishly to their private signals, which is offset by the increased size of signals received.

To circumvent this issue, I turn attention to *actual* strategic incentive $\lambda = b\delta$ faced by a voter. The typical (both modal and median) strategic incentive is then $bE[\delta] = b\eta = \xi\Phi^{-1}(\pi)$. Suppose that candidate 1 is the leading challenger ($\pi > 1/2$). Clearly, the typical strategic incentive is higher when candidate 1's position is relatively stronger. Turning back to the idiosyncrasy parameter ξ^2 , I must also address the fact that the information precision $1/\kappa^2$ may also be influenced by ξ^2 . Recall the micro-foundation in 2.3 where an individuals observes a sample of m preferences (including her own) from the electorate. This generated $\kappa^2 = \xi^2/m$ and $\text{cov}[\delta, \varepsilon] = \kappa^2 \Rightarrow \rho = \kappa/\xi = 1/\sqrt{m}$. In this environment, m represents the information available to voters. $\hat{b}(b)$ then satisfies:

$$\hat{b}(b) = \frac{2\Phi^{-1}(\gamma)\sqrt{m^2 + m(b^2 + 2b)}}{\xi(1+b)} \Rightarrow \hat{b}(0) = \frac{2\Phi^{-1}(\gamma)m}{\xi} \quad (6)$$

It is easiest to consider the decision-theoretic case. When each voter expects others to vote for their first preference and hence $b = \hat{b}(0)$ then the typical strategic incentive is:

$$\hat{b}(0)\eta = \frac{2\Phi^{-1}(\gamma)m}{\xi} \times \xi\Phi^{-1}(\pi) = 2m\Phi^{-1}(\gamma)\Phi^{-1}(\pi) \quad (7)$$

In a decision-theoretic world, then, the *net* of effect of idiosyncrasy on strategic voting incentives is zero, and the typical strategic incentive depends in a simple way on qualified majority γ , the relative strength of the leading challenger π and the amount of information m . Turning attention back b^* , idiosyncrasy continues to have an effect: increased idiosyncrasy tends to increase incentives to vote strategically. There is an additional factor present. Increased idiosyncrasy means that other voters are more less likely to vote strategically: There are fewer relatively indifferent voters. In a game-theoretic world, the caution inherent in optimal voting behavior (see Section 4.2) is reduced. Hence the incentive to vote strategically is greater. This is incorporated into Proposition 4.

Proposition 4. *Strategic voting incentives increase with the required qualified majority γ , the relative strength of the leading challenger π and the informativeness of signals m . In a stable voting*

equilibrium strategic voting is increasing in the idiosyncrasy of the electorate ξ^2 . Furthermore:

$$\frac{\hat{b}(0)\eta}{m} = 2\Phi^{-1}(\gamma)\Phi^{-1}(\pi) \quad \text{and} \quad \frac{b^*\eta}{\sqrt{m}} \longrightarrow \Phi^{-1}(\gamma)\Phi^{-1}(\pi) \sqrt{2 + 2\sqrt{\frac{(\Phi^{-1}(\gamma))^2 + \xi^2}{(\Phi^{-1}(\gamma))^2}}} \text{ as } m \rightarrow \infty$$

so that strategic voting increases with m decision-theoretically and \sqrt{m} game-theoretically.

Proof. See Appendix A. □

The last part of the Proposition 4 calculates the rate at which strategic voting changes when voters are faced with more information. The interesting feature is this: When voters are game-theoretic (so that they anticipate strategic voting by others) they react *much more slowly* to the increased availability of information. In fact, since strategic voting incentives are increasing (at least asymptotically) in \sqrt{m} , this means that the marginal effect of increased information declines to zero as $m \rightarrow \infty$. If (game-theoretic) voters are quite well informed, then any additional information sources have little effect.

Allowing $m \rightarrow \infty$ generates a benchmark case where voters are almost perfectly informed. Almost all voters receive a signal that is arbitrarily close to η , and hence face a strategic incentive $b\eta$. This strategic incentive is unboundedly large for $m \rightarrow \infty$ and hence (for $\pi > 1/2$) almost everyone successfully coordinates on candidate 1. Put simply, this means that the uniquely stable partially coordinated equilibrium is almost perfectly Duvergerian when $m \rightarrow \infty$. Palfrey's (1989) strict interpretation of Duverger's Law (almost) holds when voters have (almost) perfect knowledge of the constituency situation. Cox's (1994) classification of Duvergerian and non-Duvergerian equilibria is rejected. Only *one* fully Duvergerian outcome remains as $m \rightarrow \infty$. Interestingly, the case of $m \rightarrow \infty$ essentially retains the assumption that voters know the constituency situation, but removes the assumption that this is common knowledge. Thus the removal of the common knowledge element underpinning the Cox-Palfrey framework eliminates all but one equilibrium.

Changes in ξ^2 influence the strategic incentive faced by voters, at least in the game-theoretic case. Such changes also influence voter preferences, however, and hence potentially determine a voter's commitment to her most preferred candidate. To see this,

equip a voter with the typical signal of $\delta = \eta > 0$. She prefers candidate 2 whenever $\tilde{u} < 0$ but actually votes for candidate 1 when $\tilde{u} < 0 \leq b\eta + \tilde{u}$. The probability that an individual votes strategically, given that it makes sense for her to do so (i.e. $\delta > 0$ and $\tilde{u} < 0$) is:

$$\frac{\Pr[\tilde{u} < 0 \leq b\eta + \tilde{u} \mid \eta]}{\Pr[\tilde{u} < 0 \mid \eta]} = \frac{\Phi((1+b)\Phi^{-1}(\pi)) - \pi}{1 - \pi}$$

Of course, if all voters were to receive approximately accurate signal realizations, then the actual support received by candidate 1 would be $p = \Phi((1+b)\Phi^{-1}(\pi))$, and so $(1+b)$ is the strategic multiplier by which the support of candidate 1 is increased. From this analysis, the probability that an appropriately positioned individual (a supporter of the trailing candidate signal equipped with an accurate realization of the signal) is determined, and increasing in, b and π . Summarizing:

Proposition 5. *In the uniquely stable voting equilibrium, the probability that an appropriately positioned individual votes strategically increases with the qualified majority, the relative strength of the leading candidate and the information available to voters, but decreases with idiosyncrasy.*

Proof. Follows from the above discussion and comparative statics on b^* . □

6. CONCLUDING REMARKS

I have argued that it is critical to consider the informational underpinnings of a formal voting model. Fortunately, such considerations generate a uniquely stable voting equilibrium that exhibits partial, but not complete, strategic voting: Instrumental rationality generates only a tendency towards bipartism, in accordance with Duverger's original psychological effect. But how large is this tendency?

In companion work (Myatt 2002), I address this issue via a calibration exercise. Using data from recent British General Elections, I choose appropriate values for the constituency situation parameters π and γ . For lower levels of information (low m) the results reflect the observed incidence of strategic voting.³¹ Increasing m , however, generates rather a overly

³¹The calibration exercise also allows the potential impact of strategic voting to be calculated. It suggests that in 1997 British General Election, the Conservatives may have lost up to 60 seats from strategic voting.

large degree of strategic voting. It would appear, therefore, that when voters are particularly well-informed, non-instrumental motivations may determine vote choice. This is natural: They become increasingly certain that their vote has no effect, and hence any non-instrumental element in their preferences can take over. My conclusion is that, when voters are well informed, any formal theory must take seriously non-instrumental issues.

My final comment is that the present theory disputes the established informal intuition that strategic voting should be greater in marginal constituencies when the preferred candidate is far from contention. A reformulation of the model (Myatt 2002) demonstrates that the marginality hypothesis — that the incentive to vote strategically is greater when the margin of victory is small — is not supported:³² After controlling for the distance from contention of a preferred candidate, an expansion of the winning margin increases strategic voting incentives. Using survey data, Fisher (2000) shows that this comparative static holds in recent British elections.³³ It would appear, therefore, that a formal theory may generate predictions that are both counter-intuitive and yet consistent with observation.

APPENDIX A. OMITTED PROOFS

Proof of Lemma 2: I need to show that p takes on all values in $(0, 1)$ interval with positive density. First note that $v(\delta_i, \tilde{u}_i) = 1$ for some δ_i, \tilde{u}_i by assumption. By monotonicity, it retains this value for all larger δ_i, \tilde{u}_i . It follows that $E[v(\delta_i, \tilde{u}_i) | \eta] \rightarrow 1$ as $\eta \rightarrow \infty$. Similarly, $E[v(\delta_i, \tilde{u}_i) | \eta] \rightarrow 0$ as $\eta \rightarrow -\infty$. The posterior belief over η yields a positive density on the entire real line, and hence $f(p | \delta)$ has full support. Continuity follows straightforwardly from the properties of the expectation with respect to a continuous density. \square

Proof of Proposition 1: I introduce the parameter $\tilde{\gamma}$ where $\frac{1}{2} < \tilde{\gamma} < 1$ and define:

$$r(\tilde{\gamma}) \equiv \frac{\int_0^1 [p^{\tilde{\gamma}}(1-p)^{1-\tilde{\gamma}}]^n f(p) dp}{\int_0^1 [p^{\tilde{\gamma}}(1-p)^{1-\tilde{\gamma}}]^n dp}$$

³²Cain (1978, p. 644) provides a classic example of this hypothesis. He expects the pressure to defect (i.e. vote strategically) to be lower in “noncompetitive” constituencies.

³³Allowing uncertainty for the support of all candidates Myatt and Fisher (2002) demonstrate that a strategic incentive variable from a version of the present model is a strong predictor of strategic voting in Britain.

where $f(p)$ is a continuous density taking strictly positive values on $(0, 1)$, as per Lemma 2. Dependence on δ is suppressed for simplicity. I next introduce the notation $G(p)$:

$$G(p) \equiv \frac{p^{\tilde{\gamma}}(1-p)^{1-\tilde{\gamma}}}{\tilde{\gamma}^{\tilde{\gamma}}(1-\tilde{\gamma})^{1-\tilde{\gamma}}} \Rightarrow r(\tilde{\gamma}) = \frac{\int_0^1 G(p)^n f(p) dp}{\int_0^1 G(p)^n dp} \quad (8)$$

Notice that $G(p)$ is increasing from $G(0)$, attaining a maximum of $G(\tilde{\gamma}) = 1$ at $p = \tilde{\gamma}$, and then declining back to $G(1) = 0$. Next fix $\min\{1/4, 1 - \gamma\} > \varepsilon > 0$. For convenience, I now define $f_\varepsilon(x) = \max_{x-\varepsilon \leq p \leq x+\varepsilon} f(p) < \infty$, where the maximum is well defined since $[x - \varepsilon, x + \varepsilon]$ is a compact set and $f(p)$ is continuous from Lemma 2. I formulate an upper bound for the ratio $r(\tilde{\gamma})$ in Equation (8):

$$r(\tilde{\gamma}) \leq \frac{\bar{f}_\varepsilon(\tilde{\gamma}) \int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp + \bar{f}_{2\varepsilon}(\tilde{\gamma}) \left[\int_{\tilde{\gamma}-2\varepsilon}^{\tilde{\gamma}-\varepsilon} G(p)^n dp + \int_{\tilde{\gamma}+\varepsilon}^{\tilde{\gamma}+2\varepsilon} G(p)^n dp \right]}{\int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp} + \frac{F(\tilde{\gamma} - 2\varepsilon)G(\tilde{\gamma} - 2\varepsilon)^n + (1 - F(\tilde{\gamma} + 2\varepsilon))G(\tilde{\gamma} + 2\varepsilon)^n}{\int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp} \quad (9)$$

where obviously $F(p) = \int_0^p f(s) ds$. The right hand side of Equation (9) has five terms. I will consider them in turn. First:

$$\frac{\bar{f}_\varepsilon(\tilde{\gamma}) \int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp}{\int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp} = \bar{f}_\varepsilon(\tilde{\gamma})$$

Next, the denominator of the second term. $G(p)$ is increasing from $\tilde{\gamma} - \varepsilon$ to $\tilde{\gamma}$ and hence:

$$\int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp \geq \int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}} G(p)^n dp \geq \varepsilon G(\tilde{\gamma} - \varepsilon)^n$$

Taking the second term, and allowing $n \rightarrow \infty$, it follows that:

$$\frac{\bar{f}_{2\varepsilon}(\tilde{\gamma}) \int_{\tilde{\gamma}-2\varepsilon}^{\tilde{\gamma}-\varepsilon} G(p)^n dp}{\int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp} \leq \frac{\bar{f}_{2\varepsilon}(\tilde{\gamma})}{\varepsilon} \int_{\tilde{\gamma}-2\varepsilon}^{\tilde{\gamma}-\varepsilon} \left[\frac{G(p)}{G(\tilde{\gamma} - \varepsilon)} \right]^n dp \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

which holds since $G(p) < G(\tilde{\gamma} - \varepsilon)$ for all $p < \tilde{\gamma} - 2\varepsilon$. An identical argument ensures that the third term vanishes. A similar argument applies to the fourth term:

$$\frac{F(\tilde{\gamma} - 2\varepsilon)G(\tilde{\gamma} - 2\varepsilon)^n}{\int_{\tilde{\gamma}-\varepsilon}^{\tilde{\gamma}+\varepsilon} G(p)^n dp} \leq \frac{F(\tilde{\gamma} - 2\varepsilon)}{\varepsilon} \left[\frac{G(\tilde{\gamma} - 2\varepsilon)}{G(\tilde{\gamma} - \varepsilon)} \right]^n \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

with a symmetric argument for the fifth term. Conclude that $\lim_{n \rightarrow \infty} r(\tilde{\gamma}) \leq \bar{f}_\varepsilon(\tilde{\gamma})$. Notice now that ε may be chosen arbitrarily small. It follows that:

$$\lim_{n \rightarrow \infty} r(\tilde{\gamma}) \leq \lim_{\varepsilon \rightarrow 0} \bar{f}_\varepsilon(\tilde{\gamma}) = f(\tilde{\gamma})$$

A similar procedure bounds the limit below, and hence $r(\tilde{\gamma}) \rightarrow f(\tilde{\gamma})$. Next, I construct a compact interval $[\gamma - \psi, \gamma + \psi]$ around γ , for small ψ . For $\tilde{\gamma} \in [\gamma - \psi, \gamma + \psi]$, the argument above establishes that $r(\tilde{\gamma}) \rightarrow f(\tilde{\gamma})$ pointwise on this interval. But since $r(\tilde{\gamma})$ and its limit are continuous, and the interval is compact, it follows that this convergence is uniform. Now, recall that $\gamma_n = \lceil \gamma n \rceil / n$, and hence $\gamma_n \rightarrow \gamma$. It follows that $\gamma_n \in [\gamma - \psi, \gamma + \psi]$ for sufficiently large n . For sufficiently large n , $r(\gamma_n)$ is arbitrarily close to $f(\gamma_n)$. By taking ψ sufficiently small, it is assured that this is arbitrarily close to $f(\gamma)$, which follows from the continuity of f . Finally, to complete the proof, I note that:

$$\binom{n}{\gamma_n n} \int_0^1 [p^{\gamma_n} (1-p)^{1-\gamma_n}]^n dp = \frac{1}{n+1} \frac{\Gamma(n+2) \int_0^1 [p^{\gamma_n} (1-p)^{1-\gamma_n}]^n dp}{\Gamma(\gamma_n n + 1) \Gamma((1-\gamma_n)n + 1)} = \frac{1}{n+1}$$

This follows from spotting the density of the Beta distribution with parameters $\gamma_n n$ and $n - \gamma_n n$. It then follows that:

$$(n+1)q_1 = \frac{\int_0^1 [p^{\tilde{\gamma}} (1-p)^{1-\tilde{\gamma}}]^n f(p) dp}{\int_0^1 [p^{\tilde{\gamma}} (1-p)^{1-\tilde{\gamma}}]^n dp} = r(\gamma_n) \rightarrow f(\gamma)$$

From this the result for $(n+1)q_1$ follows, with a similar approach for $(n+1)q_2$. \square

Proof of Lemma 3: Suppose that all voters $i \neq 0$ use a symmetric and monotonic voting strategy exhibiting strict multi-candidate support. Define $p = E[v(\delta_i, \tilde{u}_i) | \eta] = H(\eta)$. This is strictly and smoothly increasing in η . Write $h(\eta) = H'(\eta)$. From the perspective of the voter $i = 0$ with signal δ , it follows that:

$$F(p | \delta) = \Pr[\eta \leq H^{-1}(p) | \delta] = \Phi \left(\frac{H^{-1}(p) - E[\eta | \delta]}{\sqrt{\text{var}[\eta | \delta]}} \right)$$

where Φ is the cumulative distribution function of the normal. This probability (via $E[\eta]$ and $\text{var}[\eta]$) is conditional on the signal δ , and uses the fact that private posterior beliefs

over η are normal. Differentiate to obtain:

$$f(p|\delta) = \frac{1}{h(H^{-1}(p))\sqrt{\text{var}[\eta|\delta]}}\phi\left(\frac{H^{-1}(p) - \mathbb{E}[\eta|\delta]}{\sqrt{\text{var}[\eta|\delta]}}\right)$$

Turning to the strategic incentive λ :

$$\begin{aligned} \log \frac{f(\gamma|\delta)}{f(1-\gamma|\delta)} &= \log \frac{h(H^{-1}(1-\gamma))}{h(H^{-1}(\gamma))} - \frac{(H^{-1}(\gamma) - \mathbb{E}[\eta|\delta])^2}{2\text{var}[\eta|\delta]} + \frac{(H^{-1}(1-\gamma) - \mathbb{E}[\eta|\delta])^2}{2\text{var}[\eta|\delta]} \\ &= \log \frac{h(H^{-1}(1-\gamma))}{h(H^{-1}(\gamma))} + \frac{(H^{-1}(1-\gamma))^2 - (H^{-1}(\gamma))^2}{2\text{var}[\eta|\delta]} + \frac{H^{-1}(\gamma) - H^{-1}(1-\gamma)}{\text{var}[\eta|\delta]}\mathbb{E}[\eta|\delta] \quad (10) \end{aligned}$$

$\mathbb{E}[\eta|\delta] = \delta$ and $\text{var}[\eta|\delta]$ does not depend on δ , and so λ is linear in δ as required. \square

Proof of Lemma 4: Under linear strategies, individual i votes for option 1 whenever $\eta + a + b\delta_i + \varepsilon_i \geq 0$, or equivalently $a + (1+b)\eta \geq -\varepsilon_i - b(\delta_i - \eta)$. Conditional on η , this last term is normally distributed with zero expectation, and variance $\tilde{\kappa}^2 = \text{var}[\varepsilon_i + b(\delta_i - \eta)]$.

Thus the probability p of a vote for candidate 1 satisfies:

$$p = H(\eta) = \Phi\left(\frac{a + (1+b)\eta}{\tilde{\kappa}}\right) \Rightarrow \eta = H^{-1}(p) = \frac{\tilde{\kappa}\Phi^{-1}(p) - a}{1+b}$$

where Φ is the cumulative distribution function of the standard normal. Differentiating:

$$h(\eta) = H'(\eta) = \frac{1+b}{\tilde{\kappa}}\phi\left(\frac{a + (1+b)\eta}{\tilde{\kappa}}\right) \Rightarrow h(H^{-1}(p)) = \frac{1+b}{\tilde{\kappa}}\phi(\Phi^{-1}(p))$$

Begin with the first term of Equation (10). First employ the symmetry of the normal distribution to note that $\Phi^{-1}(1-\gamma) = -\Phi^{-1}(\gamma)$, and that $\phi(z) = \phi(-z)$. It follows that:

$$\log \frac{h(H^{-1}(1-\gamma))}{h(H^{-1}(\gamma))} = \log \frac{\phi(\Phi^{-1}(1-\gamma))}{\phi(\Phi^{-1}(\gamma))} = 0$$

Next consider the second term of Equation (10):

$$(H^{-1}(\gamma))^2 = \frac{(\tilde{\kappa}\Phi^{-1}(\gamma) - a)^2}{(1+b)^2} = \frac{[\tilde{\kappa}\Phi^{-1}(\gamma)]^2 + a^2 - 2a\tilde{\kappa}\Phi^{-1}(\gamma)}{(1+b)^2}$$

Similarly:

$$(H^{-1}(1-\gamma))^2 = \frac{[\tilde{\kappa}\Phi^{-1}(1-\gamma)]^2 + a^2 - 2a\tilde{\kappa}\Phi^{-1}(1-\gamma)}{(1+b)^2} = \frac{[\tilde{\kappa}\Phi^{-1}(\gamma)]^2 + a^2 + 2a\tilde{\kappa}\Phi^{-1}(\gamma)}{(1+b)^2}$$

It follows that:

$$\frac{H^{-1}(1-\gamma)^2 - H^{-1}(\gamma)^2}{2\text{var}[\eta|\delta]} = \frac{2a\tilde{\kappa}\Phi^{-1}(\gamma)}{\text{var}[\eta|\delta](1+b)^2}$$

The final term is simply:

$$\frac{H^{-1}(\gamma) - H^{-1}(1-\gamma)}{\text{var}[\eta|\delta]} \mathbb{E}[\eta|\delta] = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)(1+b)}{(1+b)^2\text{var}[\eta|\delta]} \mathbb{E}[\eta|\delta]$$

Assemble these terms and, recalling that $\mathbb{E}[\eta|\delta = \delta] = \delta$ and $\text{var}[\eta|\delta] = \kappa^2$, obtain:

$$\log \frac{f(\gamma|\delta)}{f(1-\gamma|\delta)} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)(a + (1+b)\delta)}{\kappa^2(1+b)^2} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)a}{\kappa^2(1+b)^2} + \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)\delta}{\kappa^2(1+b)}$$

This is exactly the desired result. □

Proof of Lemma 5: From Lemma 4, analysis of the best response $\hat{b}(b) = 2\tilde{\kappa}\Phi^{-1}(\gamma)/\kappa^2(1+b)$ requires evaluation of $\tilde{\kappa}$, which satisfies $\tilde{\kappa}^2 = \text{var}[\varepsilon_i + b(\delta_i - \eta) | \eta] = \xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b$ where ρ is the correlation coefficient between ε_i and δ_i , conditional on the common utility component η . Hence:

$$\hat{b} = \frac{2\Phi^{-1}(\gamma)}{\kappa^2} \frac{\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b}}{1+b}$$

Evaluating the derivative:

$$\begin{aligned} \hat{b}'(b) &= \frac{2\Phi^{-1}(\gamma)}{\kappa^2} \left\{ \frac{b\kappa^2 + \rho\kappa\xi}{(1+b)\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b}} - \frac{\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b}}{(1+b)^2} \right\} \\ &= \frac{2\Phi^{-1}(\gamma)\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b}}{\kappa^2(1+b)} \left\{ \frac{b\kappa^2 + \rho\kappa\xi}{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b} - \frac{1}{1+b} \right\} \\ &= \hat{b}(b) \left\{ \frac{b\kappa^2 + \rho\kappa\xi}{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b} - \frac{1}{1+b} \right\} \end{aligned}$$

This is (weakly) decreasing for $\hat{b} \geq 0$ if:

$$\frac{b\kappa^2 + \rho\kappa\xi}{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b} \leq \frac{1}{1+b}$$

Re-arrange this expression to obtain $\xi(\rho\kappa - \xi) \leq b\kappa(\rho\xi - \kappa)$. To check this inequality, first consider the right hand side. First $\rho \geq \kappa/\xi$ by assumption — see Section 2. Since $b \geq 0$, it is sufficient to show that the left hand side is weakly negative, which requires $\xi \geq \rho\kappa$. But this holds, since $0 \leq \rho \leq 1$ and $\kappa \leq \xi$. It follows that the function is (weakly) decreasing

everywhere. Next, evaluate at the extremes to obtain:

$$\hat{b}(0) = \frac{2\xi\Phi^{-1}(\gamma)}{\kappa^2} \quad \text{and} \quad \lim_{b \rightarrow \infty} \hat{b}(b) = \frac{2\Phi^{-1}(\gamma)}{\kappa}$$

These calculations yield the desired properties of the function. \square

Proof of Lemma 6: To obtain an upper bound for the fixed point b^* , write $\hat{b}(b)$ as:

$$\hat{b}(b)\kappa = \frac{2\Phi^{-1}(\gamma)\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\xi\kappa b}}{\kappa + b\kappa}$$

Make the change of variable $\beta = \kappa b$ to obtain:

$$\hat{\beta}(\beta) = \frac{2\Phi^{-1}(\gamma)\sqrt{\xi^2 + \beta^2 + 2\rho\xi\beta}}{\kappa + \beta} \leq \frac{2\Phi^{-1}(\gamma)(\xi + \beta)}{\beta}$$

An upper bound may now be obtained by solving $\beta^2 - 2\Phi^{-1}(\gamma)(\xi + \beta) = 0$. Finding a positive root β generates an upper bound for the fixed point b^* :

$$\beta = \Phi^{-1}(\gamma) \left\{ 1 + \sqrt{\frac{\Phi^{-1}(\gamma) + 2\xi}{\Phi^{-1}(\gamma)}} \right\} \Rightarrow b^* \leq \frac{\Phi^{-1}(\gamma)}{\kappa} \left\{ 1 + \sqrt{\frac{\Phi^{-1}(\gamma) + 2\xi}{\Phi^{-1}(\gamma)}} \right\}$$

This upper bound was obtained by setting $\rho = 1$. A tighter bound is available via a formal implementation of the microfoundation for the privately observed signal. In that case, the correlation coefficient satisfied $\rho = \kappa/\xi$. The bound on the transformed equation becomes:

$$\hat{\beta}(\beta) = \frac{2\Phi^{-1}(\gamma)\sqrt{\xi^2 + \beta^2 + 2\kappa\beta}}{\kappa + \beta} = 2\Phi^{-1}(\gamma)\sqrt{\frac{\xi^2 + \beta^2 + 2\kappa\beta}{\kappa^2 + \beta^2 + 2\kappa\beta}}$$

It is clear that the right hand side is decreasing in κ . Hence sending $\kappa \downarrow 0$:

$$\hat{\beta}(\beta) \leq \frac{2\Phi^{-1}(\gamma)\sqrt{\xi^2 + \beta^2}}{\beta}$$

To obtain an upper bound, solve the equation $\beta^4 - (2\Phi^{-1}(\gamma))^2(\xi^2 + \beta^2) = 0$. This is quadratic in β^2 , and may be solved to obtain the positive root:

$$\beta^2 = \frac{(2\Phi^{-1}(\gamma))^2 + \sqrt{(2\Phi^{-1}(\gamma))^4 + 4(2\Phi^{-1}(\gamma))^2\xi^2}}{2} = \frac{(2\Phi^{-1}(\gamma))^2}{2} \left\{ 1 + \sqrt{1 + \frac{\xi^2}{(\Phi^{-1}(\gamma))^2}} \right\}$$

It follows that an upper bound is:

$$b^* \leq \frac{\Phi^{-1}(\gamma)}{\kappa} \sqrt{2 + 2\sqrt{1 + \frac{\xi^2}{(\Phi^{-1}(\gamma))^2}}}$$

Moreover, it is clear that this bound is attained as $\kappa \downarrow 0$. \square

Proof of Lemma 7: Consider the mapping $B(b) = \hat{b}^{(2)}(b) = \hat{b}(\hat{b}(b))$. Notice that $b^* = \hat{b}(b^*)$ is also a fixed point of B . Taking the derivative $B'(b) = \hat{b}'(\hat{b}(b))\hat{b}'(b)$, it follows that this is an increasing function, since $\hat{b}' \leq 0$. Consider a generic fixed point b , satisfying $B(b) = b$. Evaluate the derivative at this fixed point:

$$\begin{aligned} B'(b) &= b \left\{ \frac{\hat{b}\kappa^2 + \rho\kappa\xi}{\xi^2 + \hat{b}^2\kappa^2 + 2\rho\kappa\xi\hat{b}} - \frac{1}{1 + \hat{b}} \right\} \times \hat{b} \left\{ \frac{b\kappa^2 + \rho\kappa\xi}{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b} - \frac{1}{1 + b} \right\} \\ &= \left\{ \frac{\hat{b}^2\kappa^2 + \rho\kappa\xi\hat{b}}{\xi^2 + \hat{b}^2\kappa^2 + 2\rho\kappa\xi\hat{b}} - \frac{\hat{b}}{1 + \hat{b}} \right\} \times \left\{ \frac{b^2\kappa^2 + \rho\kappa\xi b}{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b} - \frac{b}{1 + b} \right\} \end{aligned}$$

It is clear that, for $\rho > 0$, both of these terms are less than one, and hence $B'(b) < 1$ at a fixed point. It follows that any fixed point must be a downcrossing. Further fixed points would require an upcrossing, and hence there is a unique fixed point b^* . From this it follows that $b_t \rightarrow b^*$. To see this, notice that $b_{t+2} = B(b_t)$. From the properties of B , there is the required convergence. \square

Proof of Proposition 4: The effects of γ and m follow immediately from Lemma 8. The expected strategic incentive is $b\eta = b\xi\Phi^{-1}(\pi)$ which is increasing in π . For the effect of ξ^2 :

$$b^* = \hat{b}(b^*) \quad \Rightarrow \quad \frac{db^*}{d\xi} = \frac{\partial\hat{b}(b^*)}{\partial\xi} + \hat{b}'(b^*)\frac{db^*}{d\xi} \quad \Rightarrow \quad \frac{db^*}{d\xi} = \frac{1}{1 - \hat{b}'(b^*)} \frac{\partial\hat{b}(b^*)}{\partial\xi}$$

Equation 6 yields $\partial\hat{b}(b^*)\partial\xi = -\hat{b}(b^*)/\xi = -b^*/\xi$. Combine these two expressions to obtain:

$$\frac{d[b^*\xi]}{d\xi} = b^* + \xi\frac{db^*}{d\xi} = b^* \left[1 - \frac{1}{1 - \hat{b}'(b^*)} \right] = -\frac{b^*\hat{b}'(b^*)}{1 - \hat{b}'(b^*)} > 0$$

The last inequality follows from $\hat{b}'(b^*) < 0$ (since $\hat{b}(b^*)$ is decreasing) and $\hat{b}'(b^*) > -1$ (since b^* is a stable fixed point, from Lemma 7). It remains to consider the behavior of strategic incentives as $m \rightarrow \infty$. The first equality is a restatement of Equation 7. The

limiting behavior of the equilibrium strategic incentive as $m \rightarrow \infty$ is equivalent to that when $\kappa^2 \rightarrow 0$. The proof to Lemma 6 demonstrates that κb^* attains the upper bound of Equation 4 as $\kappa \rightarrow 0$. This completes the proof. \square

REFERENCES

- CAIN, B. E. (1978): "Strategic Voting in Britain," *American Journal of Political Science*, 22(3), 639–655.
- CARLSSON, H., AND E. VAN DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.
- 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.
- DEGROOT, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York.
- DUVERGER, M. (1954): *Political Parties: Their Organization and Activity in the Modern State*. Wiley, New York.
- EVANS, G., J. CURTICE, AND P. NORRIS (1998): "New Labour, New Tactical Voting? The Causes and Consequences of Tactical Voting in the 1997 General Election," in *British Elections and Parties Review Volume 8: The 1997 General Election*, ed. by D. Denver, J. Fisher, P. Cowley, and C. Pattie, pp. 65–79. Frank Cass, London and Portland, OR.
- FEY, M. (1997): "Stability and Coordination in Duverger's Law: A Formal Model of Pre-election Polls and Strategic Voting," *American Political Science Review*, 91(1), 135–147.
- FISHER, S. D. (2000): "Intuition versus Formal Theory: Tactical Voting in England 1987–1997," Paper prepared for APSA Annual Meeting, Washington DC.
- HEATH, A., AND G. EVANS (1994): "Tactical Voting: Concepts, Measurements and Findings," *British Journal of Political Science*, 24, 557–561.
- HOFFMAN, D. T. (1982): "A Model for Strategic Voting," *SIAM Journal of Applied Mathematics*, 42, 751–761.
- JOHNSTON, R. J., AND C. J. PATTIE (1991): "Tactical Voting in Great Britain in 1983 and 1987: An Alternative Approach," *British Journal of Political Science*, 21, 95–108.
- LANOUE, D. J., AND S. BOWLER (1992): "The Sources of Tactical Voting in British Parliamentary Elections, 1983–1987," *Political Behaviour*, 14, 141–157.

- 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.
- MORRIS, S., AND H. S. SHIN (2001): "Global Games: Theory and Applications," <http://www.econ.yale.edu/~smorris/>.
- MORTON, R. H. (2001): *How Elections Rule American Politics*. Norton, forthcoming.
- MYATT, D. P. (1999): "A New Theory of Strategic Voting," Nuffield College Economics Discussion Paper.
- MYATT, D. P. (2000): "Strategic Voting under the Qualified Majority Rule," Nuffield College Economics Discussion Paper 2000-W7.
- MYATT, D. P. (2002): "Information Availability and the Impact of Strategic Voting," University of Oxford Department of Economics Discussion Paper Series no. 94.
- MYATT, D. P., AND S. D. FISHER (2002): "Everything is Uncertain and Uncertainty is Everything: Strategic Voting in Simple Plurality Elections," University of Oxford Department of Economics Discussion Paper Series.
- MYERSON, R. B., AND R. J. WEBER (1993): "A Theory of Voting Equilibria," *American Political Science Review*, 87(1), 102–114.
- NIEMI, R., G. WHITTEN, AND M. FRANKLIN (1992): "Constituency Characteristics, Individual Characteristics and Tactical Voting in the 1987 British General Election," *British Journal of Political Science*, 22, 229–240.
- 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.
- PATTIE, C., AND R. JOHNSTON (1999): "Context, Conversation and Conviction: Social Networks and Voting at the 1992 British General Election," *Political Studies*, 47, 877–889.
- RIKER, W. H. (1976): "The Number of Political Parties: A Reexamination of Duverger's Law," *Comparative Politics*, 9, 753–766.
- (1982): *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. Waveland Press, Prospect Heights, IL.
- SIMON, H. A. (1954): "Bandwagon and Underdog Effects and the Possibility of Equilibrium Predictions," *Public Opinion Quarterly*, 18, 245–253.