

ISSN 1471-0498



DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

ON THE THEORY OF STRATEGIC VOTING

David P. Myatt

Number 186

March 2004

Manor Road Building, Oxford OX1 3UQ

On the Theory of Strategic Voting

David P. Myatt

Department of Economics and St. Catherine's College, Oxford University

david.myatt@economics.ox.ac.uk

March 1, 2004¹

Abstract: Many analyses of plurality-rule elections predict the complete coordination of strategic voting, and hence support for only two candidates. Here I suggest that stable multi-candidate support will arise in equilibrium. A group of voters must partially coordinate behind one of two challenging candidates in order to dislodge a disliked incumbent. In a departure from existing models, the popular support for each challenger is uncertain. This support must be inferred from the private observation of informative signals, such as the social communication of preferences throughout the electorate, or the imperfect observation of opinion polls. The uniquely stable voting equilibrium entails only limited strategic voting and hence incomplete coordination. This is due to the surprising presence of negative feedback: an increase in the degree of strategic voting by others reduces the incentives for an individual to vote strategically. The incentive to vote strategically is lower in relatively marginal elections, after controlling for the distance from contention of a trailing preferred challenger. A calibration of the model applied to the UK General Election of 1997 is consistent with the impact of strategic voting and the reported accuracy of voters' understanding of the electoral situation. It suggests that nearly 50 seats may have been lost by the Conservative party due to strategic voting.

1. PLURALITY-RULE ELECTIONS AND STRATEGIC VOTING

1.1. **Strategic Voting and Duverger's Law.** Plurality-rule elections, where the winner is the candidate who receives the largest number of votes, are vulnerable to *strategic voting*: a voter might well switch away from her preferred candidate and toward a perceived leader, in the hope of exerting a greater influence over the outcome of the election. This phenomenon has long been the focus of both economic and political scientific study. In an influential essay, Riker (1982b) cited Droop's (1871) eloquent exposition of the strategic-voting problem:

¹I offer my grateful thanks to the many friends, colleagues, seminar participants, referees, and editors who have offered insightful comments during the long gestation of this paper, its antecedents, and companion work. Special thanks go to Steve Fisher, whose empirical analysis of strategic voting raised many of the questions that I seek to answer in this paper, and to Chris Wallace. This particular paper brings together much of the analysis previously circulated in the papers *A New Theory of Strategic Voting*, *Strategic Voting Under the Qualified Majority Rule*, and *Idiosyncrasy, Information, and the Impact of Strategic Voting*.

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

Droop’s observation incorporated two interesting ideas. First, voters may have instrumental motivations for their decisions; they may choose to vote in a way that best influences who wins the election.² Second, it is an early version of what is known to political scientists as Duverger’s law; this suggests that plurality-rule elections will tend to lead to a two-party system.³ Duverger (1954) envisaged an ongoing process in which political parties would react to the likely outcomes of plurality elections. His “psychological effect” of strategic voting was a key component of this process; however, he predicted only a *tendency* toward bipartism.

Modern theories of strategic voting offer more-dramatic predictions. Palfrey (1989) and Myerson and Weber (1993) described formal models in which third-candidate support disappears: the equilibria of these models are *strictly* Duvergerian, in the sense that all voters fully coordinate on only two candidates.⁴ Cox (1994) highlighted the existence of a second category of non-Duvergerian equilibria. They involve vote switching away from a leading candidate and toward a trailing candidate resulting in a tie for second place. Such a tie attenuates incentives to vote strategically, and hence can yield an equilibrium outcome.

It is certainly the case that, under the plurality rule, two political parties often dominate. At the level of an electoral district, however, neither the strictly Duvergerian predictions of Palfrey (1989) and Myerson and Weber (1993), nor the non-Duvergerian equilibrium emphasized by Cox (1994) serve as stylized descriptions of observed behavior. In Section 6.3 I argue that this is the case in English parliamentary constituencies. Here, however, the 1970 New York senatorial election provides a classic illustration. This election was highlighted by Riker (1982a) and others as a failure of the coordination of strategic voting.⁵ In a “three horse race” two liberal candidates, Richard L. Ottinger and Charles E. Goodell, competed

²A succinct modern definition (Fisher 2001a) is that “a tactical voter is someone who votes for a party they believe is more likely to win than their preferred party, to best influence who wins in the constituency.”

³Duverger (1954) introduced his law by noting that the “simple-majority single-ballot system favors the two-party system.” The term “simple majority” refers to the plurality rule, and the term “single ballot” indicates an environment with simple single-member districts. Riker (1982b), cited by Feddersen (1992), restated the law as asserting that “plurality election rules bring about and maintain two-party competition.”

⁴Palfrey’s (1989) summary was that “with instrumentally rational voters and fulfilled expectations, multi-candidate contests under the plurality rule should result in only two candidates getting any votes.” While here I restrict attention to the plurality rule, Myerson and Weber (1993) succeeded in characterizing equilibria for a wider variety of electoral systems, and Cox (1994) analyzed elections with multi-member districts.

⁵This example was also used, and to great effect, in the undergraduate text of Morton (2001), and Cox (1994) used it to motivate his discussion of non-Duvergerian voting equilibria.

<i>Candidate</i>	<i>Votes</i>	<i>Share</i>
James Buckley	2,288,190	39%
Charles Goodell	1,434,472	24%
Richard Ottinger	2,171,232	37%
Total	5,893,894	100%

TABLE 1. The 1970 New York Senatorial Election

against the conservative James R. Buckley.⁶ I present the outcome of this election in Table 1. A widely held belief was that the liberal vote was split between Goodell and Ottinger, allowing the win for Buckley. The outcome was certainly not Duvergerian. Neither was it non-Duvergerian in the sense of Cox (1994), since such equilibria require the challengers to tie for second place.

1.2. Coordination and Common Knowledge. The 1970 New York senatorial election, and parliamentary elections in the UK and other countries, exhibit only partial coordination of strategic voting.⁷ In this paper, I suggest that such outcomes are consistent with all but one of the assumptions made by Palfrey (1989), Myerson and Weber (1993), and Cox (1994). My point of departure is this: I suppose that there is no common knowledge of the electoral situation, and that voters must learn of this situation via the private observation of informative signals. Existing theories of voting under the plurality rule rest on the assumption that an individual votes for a particular candidate with some commonly known probability.⁸ Equivalently, the election outcome is the realization of a multinomial random variable with commonly known parameters. This assumption is fundamental to existing results; I discuss this claim here, prior to describing the new contributions of this paper.

To verify the claim, I begin by highlighting two important insights of existing theories. First, an instrumental voter cares only about situations in which two or more candidates tie for

⁶More specifically, 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 nominating Goodell as a “fusion” candidate instead supported the conservative Buckley.

⁷Analysis of strategic voting in the UK was provided by Johnston and Pattie (1991), Lanoue and Bowler (1992), and Niemi, Whitten, and Franklin (1992), *inter alios*. New research, based on British Election Study data and analyzing intuitive predictions, the bimodality hypothesis of Cox (1994) and the predictions of this paper was reported by Myatt and Fisher (2003); see also Fisher (2000, 2001b). The bipartite prediction also fails in India. Riker (1976) offered an analysis of this case, and hypothesized that reduced strategic voting was due to the presence of a clear Condorcet winner, against whom strategic voting is futile.

⁸In addition to the aforementioned papers, this is also true of more recent contribution, including papers by Myerson (1999, 2000, 2002) and Piketty (2000).

the lead; these are the only circumstances in which her vote really matters.⁹ Second, any theory of strategic voting must involve an uncertain election outcome; absent any uncertainty, almost any pattern of voting behavior yields an equilibrium.¹⁰

When the situation is common knowledge, any uncertainties in individual decisions are averaged out in a large electorate. Hence, should a tie occur, it will almost always be between the two leading candidates; furthermore, the identities of these leaders are common knowledge. For almost all voters, it is optimal to support one of these leaders: a Duvergerian outcome.¹¹ Moreover, the law of large numbers removes any real uncertainty.¹²

With this observation in hand, it is interesting to note that Droop’s (1871) logic can apply only when the identities of the parties “between whom the election really lies” are known to voters. Similarly, Cox (1997, p. 78) fully recognized that a “condition necessary to generate pure local bipartism is that the identity of trailing and front-running candidates is common knowledge.” The step taken in this paper is to remove this common knowledge, and assess the impact of voters’ information and beliefs on the strategic-voting phenomenon.¹³

1.3. Private Information and Negative Feedback. I turn now to the contributions of this paper. To focus on the central issues discussed above, I restrict to a class of three-candidate plurality-rule elections. One of the candidates, perhaps a disliked incumbent office-holder, enjoys some fixed, unwavering support. The remaining group of voters wish to dislodge this disliked incumbent.¹⁴ To do so, a proportion $\gamma > \frac{1}{2}$ of the anti-incumbent voters must vote in favor of one of two challenging candidates. This generates a *qualified majority* voting game: a qualified majority (or super-majority) γ of the dissatisfied voters must successfully coordinate if the incumbent is to be defeated. This scenario captures many

⁹An “instrumental voter” cares only about her vote insofar as it influences the winner of the election. The importance of pivotal events was, of course, fully recognized in the earlier decision-theoretic analyses of McKelvey and Ordeshook (1972) and Hoffman (1982), as well as the game-theoretic treatments cited above.

¹⁰Osborne (1995, p. 293) offered a concise summary of this point. When the outcome is known, and a voter is not pivotal (there is no tie or near-tie for the lead), instrumental concerns give no guide to her.

¹¹Cox (1994) cleverly constructed non-Duvergerian equilibria in the following way. Begin by ordering three candidates according to their popularity. From the supporters of the second candidate, a precise fraction switch away to the third candidate. This ensures that there is (almost) a tie for second place. More importantly (for the construction of an equilibrium) either of the second or third candidates may tie with the first candidate; this ambiguity prevents all support bleeding away to the leading pair.

¹²Many authors have observed that the specification of uncertainty is central to any theory of voting. Palfrey and Rosenthal (1985, p. 62), for instance, examined “the role played by uncertainty or the lack of information in affecting political processes.” They considered a voter-participation game with incomplete information over voting costs. However, the distribution of voting costs was assumed to be known, and individuals were drawn independently from it. For a large electorate, therefore, aggregate uncertainty was absent.

¹³The information and beliefs of voters concern empirical researchers: Heath and Evans (1994) criticized 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 a constituency.

¹⁴Throughout the paper, I will refer to the disliked candidate as an “incumbent.” This terminology is adopted in order to ease the exposition; the current occupations of the candidates play no part in the analysis.

elements of modern plurality elections. In the 1997 UK General Election the incumbent (and unpopular) Conservative party polled between one-third and one-half 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 candidate or the Liberal Democrat candidate in order to ensure a Tory defeat.¹⁵ This scenario is also a stylized description of the 1970 New York senatorial election (Table 1).

Amongst the anti-incumbent voters, the fraction who prefer the first challenger to the second is unknown. Voters must infer this fraction from the private observation of informative signals. Such signals might reflect either the social communication of preferences throughout the electorate, or the imperfect observation of opinion polls. The private observation ensures that the electoral situation is never common knowledge. This leads to the somewhat-surprising phenomenon of *negative feedback*: when others in the electorate are expected to vote strategically, then an individual voter faces a reduced incentive to do so.

The explanation for this negative feedback begins with an examination of the voting problem faced by an individual voter. Based on any private information available to her, such an individual will consider the relative likelihood of the different “pivotal outcomes” under which one of the challenging candidates ties with the disliked incumbent. Furthermore, the likelihood ratio of such events will be determined by her beliefs about the true underlying support for each of the challengers. If the likelihood ratio deviates away from unity, then the individual in question may face an incentive to vote strategically.

With such likelihood ratios in mind, consider a benchmark scenario in which an individual voter expects all others to vote truthfully. A relatively large bias in underlying support toward the first challenger may be required to achieve the qualified majority γ , and similarly a relatively large bias toward the second challenger may be needed for him to do the same. The voter compares the probabilities of these pivotal events which are *relatively far apart* and will have *different* probabilities of occurrence. Her incentive to vote strategically may then be large.

Suppose instead that an individual voter anticipates that others are likely to vote strategically by responding strongly to their private signals. A relatively small bias in the true underlying support toward the first challenger is all that is required for him to hit the qualified majority of γ : a small bias will lead to signals indicating the first challenger’s status as a leading contender, and hence will prompt net strategic switching away from the second challenger. This enhances the first challenger’s position, enabling him to reach the threshold γ . Identical

¹⁵Throughout the parliamentary constituencies of England, the three major political parties (Conservative, Labour, and Liberal Democrat) are the only effective competitors. The Conservative (equivalently, Tory) party is seen by many as right-wing; the other parties are seen as left-wing. Some may argue that the positions of the parties on the Downs-Black-Hotelling spectrum have changed since Labour’s rise to power. For that reason, I restrict attention to the 1997 election in Section 6.3.

logic shows that a relatively small bias in true support toward the second challenger is all that is required for him to do the same. The individual voter contemplates two situations involving *relatively smaller* biases. Such events are *relatively close* and hence will have *similar* probabilities of occurrence: the incentive to vote strategically is small.

This argument contrasts with a tale of positive feedback. According to this, more familiar, logic, strategic voting reduces the support of the less popular challenger. This loss of support (and consequent gain for the leading challenger) enhances the incentive to vote strategically, and thus further erodes the support of the trailing challenger. This is the “bandwagon effect” of Simon (1954), and leads to the elimination of any third-party support. This bandwagon logic is somewhat flawed: when voters do not have a common understanding of the electoral situation, it will not be clear who the leading challenger is. To put it somewhat crudely, a voter becomes fearful that the bandwagon is rolling in a direction opposite to that indicated by her own private signal, and is more cautious about switching her vote. Anticipation of more switching by others (a speedier bandwagon) enhances this caution.

1.4. Implications. The negative-feedback effect is somewhat subtle; nevertheless, its implications are immediate. First, it implies that game-theoretic considerations (allowing voters to anticipate strategic switching by others) actually *reduce* the impact of strategic voting. Second, negative feedback leads away from fully coordinated strategic voting behavior and toward a uniquely stable equilibrium in which there is only partial coordination: both challenging candidates enjoy positive support. Third, while the net effect of strategic voting is to help the leading challenger at the ballot box, there is also movement away from this candidate; some voters, receiving erroneous signals of the challengers’ support levels, switch in the wrong direction. Fourth, the existence of a uniquely stable equilibrium with multi-candidate support permits comparative static exercises: I am able to ascertain the implications of changes in the electoral situation; of the idiosyncrasy of voters’ preferences; and of the quality of their information sources. Furthermore, while I attempt no real empirical analysis, a calibration of the model applied to the UK General Election of 1997 resonates with observed behavior, and offers insights into the wider consequences of strategic voting. Finally, in predicting only a *tendency* toward bipartism, this paper offers results in the spirit of Duverger’s (1954) original legislation. The paper represents, therefore, not so much brand new theory, but rather a careful attempt to build a more rigorous foundation for his psychological effect.

1.5. Related Literature. Strategic voting has long been of interest to social scientists. The connections between this paper and the established literature are described throughout later sections; I offer only a few brief comments here.¹⁶ Formal studies began with decision-theoretic analyses, where voters do not take into account strategic vote-switching by others. Following Farquharson (1969), McKelvey and Ordeshook (1972) built upon work by Riker

¹⁶For a thorough review, I point to Cox’s (1997) comprehensive survey of the literature.

and Ordeshook (1968); this analysis was extended by Hoffman (1982), Cox (1984) and Niemi (1984). Later studies were game-theoretic, and examined the Nash and Bayesian-Nash equilibria of voting games. The aforementioned papers by Palfrey (1989), Myerson and Weber (1993) and Cox (1987, 1994) fall within this category. This paper builds upon all of this earlier work; as noted above my point of departure is the removal of the assumption that voters have common knowledge of the electoral situation: voters must use the information sources at their disposal to detect which strategic-voting game they are playing.

The absence of common knowledge means that the model is a global game in the sense of Carlsson and van Damme (1993); it is (Morris and Shin 2003) a game “of incomplete information whose type space is determined by the players each observing a noisy signal of the underlying state.”¹⁷ Here, the “underlying state” is the true relative popularity of the two challenging candidates. This determines the character of the game being played; in essence, the identity of the challenger best placed to beat the incumbent. The “noisy signal” corresponds to information such as social communication. In recent years, economists have made great progress by applying the techniques of global games to a wider range of problems. Applications include currency crises (Morris and Shin 1998), the pricing of debt (Morris and Shin 2004), and bank runs (Goldstein and Pauzner 2003); there are, of course, many others. There have been fewer applications to voting scenarios. Feddersen and Pesendorfer (1996, 1997, 1998) studied two-candidate models in which voters receive private information of an underlying state, and where an individual’s own payoffs depend upon the underlying state; for instance, members of a jury must decide whether to vote for the conviction of a defendant. Their main focus was upon the information-aggregation function of an election.

1.6. A Guide to the Paper. I study a finite electorate in Section 2, and specify voters’ payoffs and information sources. In Section 3, I consider optimal voting behavior in large electorates, and use my results to define an appropriate limit-game with a unit mass of voters. I find a uniquely stable equilibrium to this game in Section 4, and analyze its comparative statics in Section 5. In Section 6, I consider further micro-foundations for the information sources of voters, before applying a calibration of the model to the 1997 UK General Election.

¹⁷Both Morris and Shin (2003) and Myatt, Shin, and Wallace (2002) offered discussions of this literature.

2. STRATEGIC VOTING AS A QUALIFIED MAJORITY VOTING GAME

2.1. The Strategic Voting Problem. A plurality-rule election is deceptively simple: each member of an electorate casts a single vote; the candidate enjoying the highest vote-total wins the election.¹⁸ Despite this simplicity, there is potential for voters strategically to switch in many directions; a full analysis of strategic voting is necessarily complex.

In many situations, however, a stylized representation of the strategic-voting problem is more straightforward. In English parliamentary constituencies, for example, three major political parties compete. In 1997, the incumbent Conservative party faced challenges from the Liberal Democrat and Labour parties.¹⁹ The Conservative administration was unpopular, and hence many voters wished to see the defeat of Tory candidates. Such voters faced the following problem: should they vote for their favorite candidate, or switch to a second choice? A switch would make sense if a vote for the second choice was more likely to stop a Conservative win. This characterization of the 1997 election suggests an analysis of two-way switching between supporters of the Labour and Liberal Democrat parties.

2.2. Qualified-Majority Voting. To model this stylized game of “beat the incumbent” I begin by fixing the votes cast for an otherwise-disliked incumbent candidate $j = 0$ at x_0 ; the incumbent will win unless another candidate obtains more votes than this. A group of $n + 1$ anti-incumbent voters is indexed by $i \in \{0, 1, \dots, n\}$. Each individual i casts a single vote for one of two challenging candidates $j \in \{1, 2\}$; the votes cast for challenger j total x_j , so that $x_1 + x_2 = n + 1$. Next, I assume that the votes cast for the incumbent satisfy

$$x_0 = \lceil \gamma n \rceil, \quad \text{where} \quad \frac{1}{2} < \gamma \leq 1,$$

and where $\lceil \gamma n \rceil$ indicates the least integer that is weakly greater than γn . The restriction $\gamma > \frac{1}{2}$ ensures that a *qualified majority* (or super-majority) of anti-incumbent voters must support one of the challengers in order to ensure the incumbent’s defeat; this generates an environment in which strategic voting might take place, and further implies that the incumbent candidate can never trail in third place.²⁰ The restriction $\gamma \leq 1$ ensures that sufficient coordination by anti-incumbent voters (for instance, when $x_j = n + 1$ for some $j \in \{1, 2\}$) will prevent an incumbent win; γ indexes the exact degree of coordination

¹⁸There may, of course, be a tie for the lead, and it is the (perhaps remote) possibility of such ties that will be of interest to any instrumentally motivated voter. In UK parliamentary elections, the returning officer enjoys a casting vote in the event of a tie, and breaks this tie with the aid of a fair coin.

¹⁹Prior to 1997, the Tories held overall power. They did not, however, hold every constituency. As in Section 1, the label “incumbent” is used in a loose fashion; the candidate so labelled is simply the one with fixed support, and against whom other voters wish to coordinate.

²⁰A qualified majority (equivalently, Black’s (1948) special majority) may be used to eliminate Condorcet cycles; see Simpson (1969), Kramer (1973, 1977), Greenberg (1979), and Caplin and Nalebuff (1988).

required. Turning to the election outcome, and based upon the vote totals, the winner is

$$j = \begin{cases} 0 & \text{if } \max\{x_1, x_2\} \leq x_0, \\ 1 & \text{if } x_1 > x_0, \text{ and} \\ 2 & \text{if } x_2 > x_0. \end{cases} .$$

Notice that the incumbent wins whenever $x_0 = x_j$ for $j \in \{1, 2\}$. This is without loss of generality; nothing would change if the winner were randomly chosen in the event of a tie.

An inspection of Table 1 reveals that the 1970 New York senatorial election fits within this story: Buckley was the disliked incumbent, with $x_0 = 2,288,190$ votes. Votes cast for Goodell and Ottinger, corresponding to anti-incumbent liberals, totalled $n + 1 = 3,605,704$. The critical qualified-majority needed for a conservative defeat was then $\gamma = x_0/n \approx 0.635$.

The assumptions that the incumbent candidate enjoys a fixed number of votes and that the remaining $n + 1$ voters are “anti-incumbent” means that incumbent supporters play no part in the analysis that follows. For this reason, all references to voters, the constituency, and the electorate, will indicate the $n + 1$ anti-incumbents.

2.3. Preferences. Payoffs are contingent only upon the identity of the winning candidate: voter i receives a payoff u_{ij} when j wins the election. All $n + 1$ voters strictly prefer both challengers to the incumbent. This permits a normalization of $u_{i0} = 0$, so that $\min\{u_{i1}, u_{i2}\} > 0$. Later analysis will confirm that the ratio of the payoffs u_{i1} and u_{i2} is sufficient to describe a voter’s preferences; taking logarithms, I define $\tilde{u}_i = \log[u_{i1}/u_{i2}]$, so that the sign of \tilde{u}_i determines the identity of voter i ’s preferred challenger, and the size $|\tilde{u}_i|$ determines the intensity of this first preference. The log relative-preference \tilde{u}_i is broken down into two components.

Assumption 1. \tilde{u}_i is decomposed into common and idiosyncratic components,

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

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

$\eta = E[\tilde{u}_i]$ is the expected log relative-preference across the population, conditional on the electoral situation. One interpretation is that η represents common factors affecting all voters, whereas ε_i represents the idiosyncratic preference of voter i . A second is that η is the log relative-preference, and hence essential identity, of the median anti-incumbent voter.

Following Assumption 1, the electoral situation may be described by the parameters γ , η and ξ . Alternatively, η may be inferred from the fraction π of the electorate who prefer the first challenger: $\pi = \Pr[\tilde{u}_i > 0] = \Phi(\eta/\xi)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. This may be inverted to obtain $\eta = \xi\Phi^{-1}(\pi)$. Finally, the specification for the distribution of ε_i permits a simple underlying foundation for the

signal technology developed below. It also offers the following convenience: ξ^2 provides a measure of the degree of preference idiosyncrasy throughout the electorate.

2.4. Information and Social Communication. Voter i is assumed to know \tilde{u}_i , but not its decomposition into η and ε_i .²¹ This means that the exact identity of the median anti-incumbent voter η (and hence the proportion $\pi = \Phi(\eta/\xi)$ who prefer the first challenger) is unknown; it must be inferred from any information sources available to a voter.²²

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.²³ Tighter beliefs are generated following the acquisition of information pertaining to this situation. This information is encapsulated in an informative signal $\delta_i | \eta \sim N(\eta, \kappa^2)$ of the common component η , with associated precision $1/\kappa^2$. Before stating this specification as Assumption 2, I first consider two examples of the information sources that might be available to a voter.

A first source of information is this: a voter may employ \tilde{u}_i as an estimate of η .²⁴ This yields a signal $\delta_i = \tilde{u}_i$ with variance $\kappa^2 = \xi^2$. Moving beyond simple introspection, a more precise signal may be obtained via a second source: the social communication of preferences throughout the electorate.²⁵ To explore this idea, suppose that voter i observes the preferences of $m - 1$ randomly chosen members of the electorate, indexed by k .²⁶ Augmenting these observations with her own log relative-preference \tilde{u}_i , this sample (of total size m) allows her to estimate η .²⁷ Given the normality assumption, the sample mean is a sufficient statistic

²¹This contrasts with the models of Feddersen and Pesendorfer (1996, 1997, 1998), in which “voters are uncertain about the realization of a state variable that affects the utility of all voters.” The distinction is that, in their work, the voter’s own payoff is unknown; this leads to their “swing voter’s curse.”

²²Qualified-majority voting is thus a global game (Carlsson and van Damme 1993); recall (Section 1) that Morris and Shin’s (2003) definition is a game “of incomplete information whose type space is determined by the players each observing a noisy signal of the underlying state.” The state variable in this case is η , and the noisy signal of the underlying state is given in Assumption 2. See also Myatt, Shin, and Wallace (2002).

²³An 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 of Assumption 2.

²⁴In a recent contribution Goeree and Grosser (2003, p. 2) noted that “people tend to use their own tastes and beliefs as information in guessing what others like and believe.”

²⁵Pattie and Johnston (1999) demonstrated that the contextual effects of conversations with family, acquaintances and others were associated with vote-switching behavior in the UK General Election of 1992.

²⁶This differs from the bounded-rationality sampling equilibrium described by Osborne and Rubinstein (2003), in which voters observe a sample of the actions taken by others.

²⁷I am implicitly assuming that she can elicit the *true* preferences of those within her sample rather than their *stated* preferences; sampled individuals might choose to misrepresent their preferences in order to strategically manipulate her beliefs. Of course, the voter 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 to hide their true colors.

for η , generating an aggregate signal

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

Hence an informative signal, with precision m/ξ^2 , arises from a detailed “private opinion poll” of size m . This size would correspond to the number of people with whom an individual interacts, so long as such people are drawn at random from the population.

Notice that the inclusion of \tilde{u}_i within the sample mean of a voter’s social communication ensures that the her posterior beliefs depend only on the sufficient statistic δ_i . I state the formal specification of a voter’s signal in terms of such a sufficient statistic.

Assumption 2. *Voter i privately observes a signal $\delta_i | \eta \sim N(\eta, \kappa^2)$. Conditional on η , the private signal δ_i and idiosyncratic component ε_i are joint normal. The signal encapsulates all payoff-relevant information for voter i ; formally it is a sufficient statistic for η .*

The restriction of the signal to a single dimension is without loss of generality. If a voter were to observe a multi-dimensional normally distributed signal then δ_i would be the appropriate weighted average of the signal components. Furthermore, the fact that a voter’s own log relative preference is informative ensures that, conditional on η , \tilde{u}_i and δ_i are correlated. In the introspection case, $\delta_i = \tilde{u}_i$, and hence $\text{cov}[\delta_i, \tilde{u}_i | \eta] = \xi^2 = \kappa^2$. Similarly, for the social-communication case, it is straightforward to observe that $\text{cov}[\delta_i, \tilde{u}_i | \eta] = \xi^2/m = \kappa^2$. In fact, δ_i ’s status as a sufficient statistic for η ensures that the covariance is always equal to the variance of the signal, as the following lemma confirms.²⁸

Lemma 1. *Conditional on η , a voter’s signal δ_i and log relative-preference \tilde{u}_i are positively correlated: $\text{cov}[\delta_i, \tilde{u}_i | \eta] = \kappa^2$, yielding a corresponding correlation-coefficient of $\rho = \kappa/\xi$.*

Following her observation of δ_i , voter i updates her diffuse prior to form a posterior belief $\eta | \delta_i \sim N(\delta_i, \kappa^2)$.²⁹ $\delta_i = E[\eta | \delta_i]$ is voter i ’s expectation of the median voter’s identity and $1/\kappa^2$ is the precision of her posterior beliefs. Importantly, different voters observe different signal realizations, and hence entertain different beliefs about the likely support for each of the challengers; κ^2 measures the variation in such opinions across the electorate.

2.5. Signal Accuracy. A voter can always rely on introspection, and hence her signal precision will be at least $1/\xi^2$. Within the context of social-communication, the precision of a voter’s posterior beliefs satisfies $1/\kappa^2 = m/\xi^2$: it increases linearly with m . This suggests

²⁸Proofs of any formal claims are contained in the appendices.

²⁹Alternatively, if she begins with the prior $\eta \sim N(\mu, \sigma^2)$, then 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 $\eta | \delta_i \sim N(\delta_i, \kappa^2)$.

that the accuracy of a voter’s perception of the electoral situation will grow with the size of her social network. If voters sample individuals who are similar to them, however, then the effective precision of their information (as measured by m) may be dramatically lower.

Suppose, for instance, that individuals live within communities, and interact only with members of their own community, where a community is a small subset of the electorate. 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 $m \leq 5$. I explore this idea in Section 6 below.

Naturally, a voter’s information is not limited to social communication. A further source might be the publication of opinion polls or other media input.³⁰ In many elections, however, these tend to occur at the *national* level, whereas candidates are elected at a *regional* level. At a regional level, opinion polls are less common.³¹

Nevertheless, the possibility of opinion polls must be taken seriously, and may still be viewed within the context of Assumption 2. For instance, if a poll were to perfectly identify η , then κ^2 would correspond to any (possibly small) noise in a voter’s observation of it. 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.³² Allowing $\kappa^2 \rightarrow 0$ (or equivalently $\alpha \rightarrow 1$) generates the perfect observation of a perfectly revealing public information source.

Returning to the social-communication micro-foundation described in Section 2.4, the notion of a signal’s accuracy may also be used, and α may be employed as a primitive parameter in the specification of the voter’s information. Since $\eta = \xi\Phi^{-1}(\pi)$, and $\kappa^2 = \xi^2/m$, then $\alpha = \Phi(\sqrt{m}|\Phi^{-1}(\pi)|)$, or equivalently $m = [\Phi^{-1}(\alpha)/\Phi^{-1}(\pi)]^2$. 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 observation of a private sample of approximately 17 individuals. Put differently, a voter with a social network of 17 people will identify the wrong candidate as the leading challenger with 15% probability. It is worth emphasizing that this statement continues to lean upon the assumption that a voter’s social network forms a representative sample of the wider anti-incumbent society.

³⁰McKelvey and Ordeshook (1985) explored a model in which voters base their decision on such sources.

³¹The 1997 UK General Election provides an example. Evans, Curtice and Norris (1998) noted 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.

³²Suppose, for instance, that $\eta > 0$, so that candidate $j = 1$ is the leading challenger. The signal δ_i correctly identifies him when $\delta_i > 0$. Thus the accuracy of the signal is $\alpha = \Pr[\delta_i > 0 | \eta] = \Phi(\eta/\kappa)$.

3. OPTIMAL VOTING BEHAVIOR IN LARGE ELECTORATES

In Section 4, I will consider a voting game in which the electorate of anti-incumbent voters is a unit mass. For such a game, payoffs are not immediately defined: an individual voter has no effect on the election outcome. To specify payoffs for such a game, I must first consider a voter's optimal behavior in a finite electorate. Taking an appropriate limit as $n \rightarrow \infty$ then yields an analytically tractable game played by a continuum of voters.

3.1. Optimal Voting Behavior. Consider the decision of voter $i = 0$; on occasion, I will refer to her as the focal voter, and for notational simplicity I will drop the subscript i when I consider her decisions. She may only influence the election when she is *pivotal*. This happens if, absent her vote, there is a tie between a challenger and the disliked incumbent. To describe pivotal events, I use x to denote the total number of votes cast for candidate 1 among the n other anti-incumbent voters.³³ If $x = x_0 = \lceil \gamma n \rceil$, then one more vote will allow candidate 1 to win. Similarly, if $n - x = x_0 = \lceil \gamma n \rceil \Leftrightarrow x = \lfloor (1 - \gamma)n \rfloor$ then candidate 2 is in the same position.³⁴ 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

$$q_1 = \Pr[x = \lceil \gamma n \rceil] \quad \text{and} \quad q_2 = \Pr[x = \lfloor (1 - \gamma)n \rfloor],$$

where I recall that $x_0 = \lceil \gamma n \rceil$ is the (fixed) vote total for the disliked incumbent. A vote for candidate 1 will generate a payoff gain of u_1 with probability q_1 , and hence an expected payoff gain of $q_1 u_1$. Similarly, a vote for candidate 2 yields an expected payoff gain of $q_2 u_2$. These observations lead to the following simple lemma.

Lemma 2. *For voter $i = 0$ with payoffs u_1 and u_2 , an optimal voting rule must satisfy:*

$$\text{vote for } \begin{cases} j = 1 & \text{if } q_1 u_1 > q_2 u_2, \\ j = 2 & \text{if } q_2 u_2 > q_1 u_1, \text{ and} \\ j = 1 \text{ or } j = 2 & \text{if } q_1 u_1 = q_2 u_2. \end{cases}$$

Notice that whenever a challenger enjoys strong support (that is, whenever $x > \lceil \gamma n \rceil$ or $1 - x > \lceil \gamma n \rceil$) 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 γ from within the population of anti-incumbent voters. When this is possible for both candidates (so that $\min\{q_1, q_2\} > 0$) the following definition may be employed.

Definition 1. *The pivotal log likelihood-ratio or strategic incentive is $\lambda \equiv \log[q_1/q_2]$.*

³³Hence, if the focal voter $i = 0$ votes for candidate 1, then $x_1 = x + 1$ and $x_2 = n - x$.

³⁴Here $\lfloor y \rfloor$ indicates the greatest integer that is weakly less than y .

Using this definition, together with Lemma 2, voter $i = 0$ should vote for $j = 1$ whenever

$$\frac{u_1 q_1}{u_2 q_2} > 1 \quad \Leftrightarrow \quad \log \left[\frac{u_1}{u_2} \right] + \log \left[\frac{q_1}{q_2} \right] > 0 \quad \Leftrightarrow \quad \tilde{u} + \lambda > 0.$$

She balances her *relative preference* for the challenging candidates (represented by \tilde{u}) against the *relative likelihood* that her vote influences the outcome of the election (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, and hence she faces no strategic incentive; she votes straightforwardly for her most preferred candidate.

Lemma 2 describes optimal behavior contingent on a voter's beliefs about pivotal events. She will use her signal and her expectation of the strategies used by others to form these beliefs. At this point I restrict attention to identity-independent strategies that are contingent only on payoff-relevant information: according to the following definition, a voter's choice depends only upon the realization of \tilde{u}_i and δ_i , and not on the index i .³⁵

Definition 2. *An identity-independent 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. This strategy is monotonic if it is (weakly) increasing in its arguments δ_i and \tilde{u}_i . It involves strict multi-candidate support if it takes both of the values 0 and 1 for appropriate choices of δ_i and \tilde{u}_i . Finally, it is fully coordinated if either $v(\delta_i, \tilde{u}_i) = 0$ for all δ_i, \tilde{u}_i or $v(\delta_i, \tilde{u}_i) = 1$ for all δ_i, \tilde{u}_i .*

Notice that the specification of an identity-independent strategy yields a strategy profile for all voters. A monotonic strategy means that an increase in the preference and signal for a challenger cannot reduce the probability of a vote for that candidate. A voting strategy exhibits multi-candidate support if a voter supports each candidate with certainty, for an appropriate choice of signal and preferences. Finally, a fully coordinated voting-strategy involves a definite vote for one of the candidates independent of the signal and preference realization.

3.2. No Electoral Uncertainty. What would happen if the common component η (and hence the identity of the median anti-incumbent voter, which also corresponds to the true underlying support π for the first challenger) were known? Restricting to the identity-independent strategies of Definition 2, votes are contingent solely on realized payoffs and signals. Conditional on η , these payoffs and signals are independently distributed. It follows that, again conditional on η , voting decisions are independent, and I may write

³⁵In this sense, strategy profiles that satisfy Definition 2 are symmetric. This does not, however, mean that all voters take the same action: a realized vote depends upon individual preference and signal realizations. A desire to maintain this distinction motivates the (slightly more awkward) identity-independent qualifier.

$p = [v(\delta_i, \tilde{u}_i) | \eta]$ for the (statistically independent) probability that a randomly selected individual votes for candidate 1.³⁶ The vote total x for candidate 1 among the n individuals $i \geq 1$ follows a binomial distribution with parameters p and n . The pivotal probability of a tie involving challenger $j = 1$ satisfies

$$q_1 = \Pr[x = \lceil \gamma n \rceil] = \binom{n}{\lceil \gamma n \rceil} p^{\lceil \gamma n \rceil} (1-p)^{\lfloor (1-\gamma)n \rfloor} \longrightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and similarly for q_2 : in a large electorate, the *absolute* probability of a pivotal outcome falls to zero. As is well understood, however, it is not the absolute but rather the *relative* likelihood of pivotal events, captured by the strategic incentive λ , that determines voting behavior. Taking the position of the voter $i = 0$, λ may be evaluated. Defining $\gamma_n = \lceil \gamma n \rceil / n$,³⁷

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

Unless $p = 1/2$, the size of the strategic incentive λ grows without bound as the electorate grows large.³⁸ 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 would be complete coordination, with the entire electorate strategically abandoning candidate 2.

3.3. Uncertain Common Effect. But what if η (the median anti-incumbent voter's identity) 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 strategies (Definition 2). If the strategy adopted by others is fully coordinated, then irrespective of η , there is no possibility of a pivotal outcome.³⁹ Consider monotonic 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 the focal voter's conditional beliefs over p .

Lemma 3. *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 $(0, 1)$.*

With this in hand, the pivotal probability of a tie involving candidate $j = 1$ becomes

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

³⁶Notice that p is the probability of a *vote* for candidate 1, whereas $\pi = \Pr[\tilde{u}_i \geq 0 | \eta]$ is the probability that a voter's *first preference* is for candidate 1. In the absence of any strategic voting, $p = \pi$.

³⁷This notation helps me to deal with integer issues in a convenient way; note that $\gamma_n \rightarrow \gamma$ as $n \rightarrow \infty$.

³⁸Employing the terminology of Myerson (2002), only one of the challenging candidates is "serious" whereas the other is "out of contention."

³⁹Thus q_1 and q_2 are exactly zero, rather than zero in the limit as $n \rightarrow \infty$. In other words, if a voter expects all others to fully coordinate, then strategic considerations will offer no guide to her vote choice.

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 individual is unknown, since it rests upon the realization of her signal and preferences. The second term $f(p|\delta)$ represents *electoral* uncertainty: from the perspective of voter $i = 0$, the electorate-wide support for candidate 1 (represented by p) is unknown. For models in which the electoral situation is common knowledge, idiosyncratic uncertainty is present, but electoral uncertainty is not.

As $n \rightarrow \infty$, the probability q_1 vanishes to zero. Its asymptotic properties are interesting, however; they are recorded in the following result, which represents a minor extension of Proposition 1 from Chamberlain and Rothschild (1981).⁴⁰

Proposition 1. *Under the conditions of Lemma 3, the pivotal probabilities satisfy*

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

This is an application of the law of large numbers: in a large electorate, the relative likelihood of the two pivotal events is the relative likelihood that the challengers' support levels (represented by p and $1-p$) coincide with the critical qualified-majority γ . Importantly, then, it is only electoral uncertainty over p (generated by uncertainty over the common effect η) that matters; equivalently, when the electoral situation is not common knowledge, then idiosyncratic uncertainty plays no role. Why is this? As the electorate grows large, the idiosyncrasies 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.⁴¹

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 open the possibility that strategic voting may be self-reinforcing, leading to an *equilibrium* outcome that involves full coordination of strategic voting—an issue that is addressed in Section 4.

3.4. Voting in Large Electorates. My next step is to analyze a voting game in a large electorate (by allowing $n \rightarrow \infty$). 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 response to any vote vanish to zero as the electorate grows large. I have observed, however, that it is not the *absolute* values of payoffs and probabilities that are of importance, but rather their *relative* size. In the finite-population voting-game, I may

⁴⁰Chamberlain and Rothschild (1981) were primarily concerned with the order of the probability of a tied outcome as $n \rightarrow \infty$; they noted that “[T]here have been to our knowledge few attempts to calculate the relevant probability. The calculations which we know of are neither general nor rigorous.” Their result was used by Mulligan and Hunter (2003) in their assessment of the empirical frequency of a pivotal vote.

⁴¹This is not a new observation; Gelman, Katz, and Tuerlinckx (2002, p. 429) explained that idiosyncratic uncertainty is removed “because the binomial variation from which it derives is minor compared to any realistic variation among the probabilities.” In other words, only electoral uncertainty matters. They noted that this observation was also made by Good and Mayer (1975) and Margolis (1983).

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 a “limit game” (with a unit mass of voters) may then be defined.

I restrict attention to the focal voter’s payoffs whenever all other voters $i \neq 0$ adopt a monotonic voting strategy with multi-candidate support; this will be sufficient for the analysis of Section 4. I observe that $(n + 1)q_1 \rightarrow f(\gamma | \delta)$ and $(n + 1)q_2 \rightarrow f(1 - \gamma | \delta)$ as $n \rightarrow \infty$, via Proposition 1. This suggests a payoff of $u_1 f(\gamma | \delta)$ in response to a vote for $j = 1$, and a corresponding payoff of $u_2 f(1 - \gamma | \delta)$ in response to a vote for $j = 2$.

This discussion leads me to specify the limit game in the following way. There are a continuum of anti-incumbent voters, indexed by $i \in [0, 1]$. Each voter i chooses a vote $j \in \{1, 2\}$, conditional on the realization of \tilde{u}_i and δ_i , which in turn follow the specifications described in Section 2. It remains for me to specify an individual voter’s payoffs.

Assumption 3. *Consider a limit game, with a unit mass of voters. When voters use a monotonic strategy with multi-candidate support (Definition 2) an individual’s payoffs satisfy*

$$\begin{aligned} U(\text{Vote } 1 | v(\delta_i, \tilde{u}_i), \delta) &\equiv f(\gamma | \delta) u_1, \quad \text{and} \\ U(\text{Vote } 2 | v(\delta_i, \tilde{u}_i), \delta) &\equiv f(1 - \gamma | \delta) u_2. \end{aligned}$$

In other circumstances, payoffs may be specified arbitrarily.

3.5. The Wasted-Vote Phenomenon. Re-scaling the payoffs by the electorate size may be an appropriate specification of payoffs. A familiar critique of voting models is that all votes are wasted in a large electorate.⁴² Proposition 1 confirms that $q_j \rightarrow 0$ at a rate of $1/(n + 1)$ as $n \rightarrow \infty$.⁴³ An expansion of the electorate, however, may also change the size of the payoffs received by a voter in the event of a win by each of the challengers. If the electorate size $n + 1$ were to measure the importance of the election in the eyes of a voter, then $(n + 1)u_j$ might well be an appropriate specification of the payoff when j wins. In other words: voters may care more about more important elections. The presence of electoral uncertainty (in the form of the density $f(p)$) ensures that such scaled payoffs are well-defined in the limit.

⁴²Setting forth his views in the form of a dialogue, Meehl (1977) argued that “all economic theories of voter participation are radically incoherent, because such participation is irrational as an instrumental action toward an egocentric end.” Similarly, Palfrey and Rosenthal (1985) asked: “why does anyone bother to vote, given that voting is presumably costly and that the probability that one’s vote will affect the outcome is presumably small?” Responses to this question have been provided by Ferejohn and Fiorina (1974), Palfrey and Rosenthal (1983), Ledyard (1984), Morton (1991), Feddersen (1992), and Aldrich (1993), *inter alios*.

⁴³The observation that the probability of a tie, when p is unknown, is of order $1/n$ is due to Chamberlain and Rothschild (1981). When p is known, then the election outcome is realization of a binomial random variable. For $p = \frac{1}{2}$, the probability of a tie is of order $1/\sqrt{n}$, and when $p \neq \frac{1}{2}$ it is of order $\exp(-cn)$ for some $c > 0$; see also Beck (1975) and Margolis (1977).

4. OPTIMAL VOTING AND EQUILIBRIUM

I now focus exclusively on the limit game of Section 3.4, corresponding to $n \rightarrow \infty$, and envisage a unit mass of voters with payoffs as specified in Assumption 3.

4.1. Linear Voting Strategies. I begin with the optimal behavior of an individual voter when all others adopt a monotonic voting strategy exhibiting multi-candidate support.

Lemma 4. *If all other voters adopt a monotonic voting strategy exhibiting strict multi-candidate support (Definition 2), then the best response for an individual is to use a linear voting strategy: she votes for candidate 1 whenever $\tilde{u} + a + b\delta \geq 0$, for some $a \in \mathbb{R}$, $b \in \mathbb{R}_+$.*

Lemma 4 ensures that, so long as voters react positively to both their payoffs and their private information, then I may focus on the class of simple linear strategies. This is a consequence of the following reasoning: when members of the electorate employ a voting strategy that is positively related to signals and preferences, then there will be two values η_1 and $\eta_2 < \eta_1$ of the common component that result in $p = \gamma$ and $p = 1 - \gamma$ respectively. An individual 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. This leads to a strategic incentive that satisfies

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

for some a and $b > 0$. An individual voter finds it optimal to vote for candidate 1 whenever $\tilde{u} + \lambda(\delta) \geq 0$,⁴⁴ which is equivalent to $\tilde{u} + a + b\delta \geq 0$.

The linearity of optimal voting strategies yields these implications: first, the class of identity-independent monotonic strategies exhibiting multi-candidate support is closed under best response, since the linear strategies of Lemma 4 are within this class; second, the class of linear strategies is closed under best response; third, any identity-independent 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 5. *If voters adopt a linear strategy $v(\delta_i, \tilde{u}_i) = 1 \Leftrightarrow \tilde{u}_i + a + b\delta_i \geq 0$, then the best response for an individual is to adopt a linear strategy $v(\delta_i, \tilde{u}_i) = 1 \Leftrightarrow \tilde{u} + \hat{a} + \hat{b}\delta \geq 0$, with*

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

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

⁴⁴Without loss of generality, I assume that she votes for candidate 1 when indifferent.

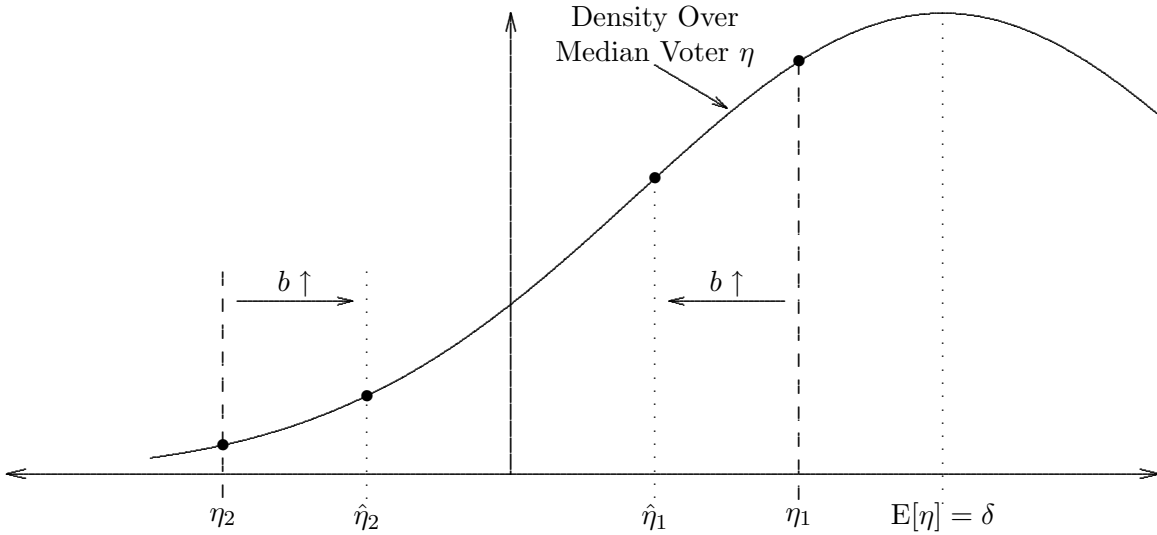


FIGURE 1. Explaining Negative Feedback

Linear strategies are easily interpreted. The intercept a is the strategic incentive faced by a voter following 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 5 a voter will exhibit a bias (e.g. $\hat{a} > 0$) if and only if she expects other voters 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 an individual's use of her signal changes in response to increased ($b \uparrow$) or decreased ($b \downarrow$) responsiveness on the part of others; an issue to which I now turn.

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 (I demonstrate this fact, just below) the appropriate vehicle for evaluating this hypothesis is b .

Lemma 5 states that $\hat{b}(b)$ is decreasing: an *increase* in the tendency by others to vote strategically ($b \uparrow$) *reduces* the tendency of an individual to do so ($\hat{b} \downarrow$). Figure 1 helps me to explain this phenomenon. Suppose that all voters cast in favor of their true first preference (equivalent to $a = 0$ and $b = 0$) and that an individual 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 from within the anti-incumbent electorate 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 $\eta = \eta_1 = \xi\Phi^{-1}(\gamma)$. Similarly, candidate 2 will just reach the qualified majority when $\eta = \eta_2 = -\xi\Phi^{-1}(\gamma)$. The focal voter will compare the relative likelihood of the events $\eta = \eta_1$ and $\eta = \eta_2$. From Figure 1, $\eta = \eta_1$ is rather more likely, and hence she faces a large incentive to vote strategically for candidate 1.

Suppose instead that all voters are willing to vote strategically by responding strongly to their signals, so that $b > 0$. The support for candidate 1 just reaches γ when $\eta = \hat{\eta}_1$, where

$$\begin{aligned} \gamma = \Pr[\tilde{u}_i + b\delta_i \geq 0 \mid \eta = \hat{\eta}_1] &= \Phi\left(\frac{(1+b)\hat{\eta}_1}{\sqrt{\text{var}[\tilde{u}_i + b\delta_i \mid \eta = \hat{\eta}_1]}}\right) = \Phi\left(\frac{(1+b)\eta}{\sqrt{\xi^2 + (b^2 + 2b)\kappa^2}}\right) \\ &\Rightarrow \hat{\eta}_1 = \frac{\Phi^{-1}(\gamma)\sqrt{\xi^2 + (b^2 + 2b)\kappa^2}}{1+b}. \end{aligned}$$

where the evaluation of the variance term employs Lemma 1. For the common component to reach $\hat{\eta}_1$ is a less extreme criterion than for it to reach η_1 , since

$$\hat{\eta}_1 = \frac{\Phi^{-1}(\gamma)\sqrt{\xi^2 + (b^2 + 2b)\kappa^2}}{1+b} < \xi\Phi^{-1}(\gamma) = \eta_1,$$

where the inequality follows from simple algebra and $\kappa^2 \leq \xi^2$. By the same logic, candidate 2 will just reach γ when $\eta = \hat{\eta}_2 = -\hat{\eta}_1$. An individual voter 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 challenging candidate 1 is much more likely to contend for the qualified majority, and this reduces the incentive to vote strategically.

Although negative feedback may seem surprising, it accords with intuition upon further reflection. When b is high, individuals respond strongly to their signals. This increases the probability of a strategic vote in *both* directions. A voter with signal $\delta > 0$ is concerned that others 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 an individual has no effect; she can *only* influence the outcome when there is a tie. Now, if there were such 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 satisfying $\eta < 0$.

4.3. Voting Equilibria with Partial Coordination. I seek monotonic equilibria exhibiting strict multi-candidate support. Following Lemma 4, these must involve linear strategies: 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 then it is a best response for a voter to use a linear strategy with coefficients $\hat{a}(a, b)$ and $\hat{b}(b)$ (Lemma 5). For an equilibrium, I require a pair $\{a^*, b^*\}$ satisfying $a^* = \hat{a}(a^*, b^*)$ and $b^* = \hat{b}(b^*)$. The second equation involves only b and hence may be considered independently. Lemma 7, contained in Appendix A, confirms that this equation has a unique solution, and obtains bounds for it. Turning attention back to a^* , the fixed point equation $a^* = b^*a^*/(1 + b^*)$ has a unique solution at $a^* = 0$.

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

Hence the partial coordination of strategic voting (implying multi-candidate support) is consistent with equilibrium behavior on the part of instrumental voters.

Other 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 electoral situation, as measured by $b^* > 0$. Of course, this is a response to a voter's *signal* of the electoral 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.⁴⁵

4.4. Equilibrium Selection. The partially coordinated equilibrium described in Proposition 2 is unique in the class of monotonic equilibria with multi-candidate support. Moving beyond this class, there may be (depending critically on a fuller specification of payoffs beyond those made in Assumption 3) *strictly* Duvergerian equilibria: 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 and hence all instrumental voters will be exactly indifferent between the two candidates. Following a heavy exploitation of any such indifference, it is an equilibrium for them to all coordinate on a single candidate. Such strictly Duvergerian equilibria are, in the context of the present model, unsatisfactory and for a number of reasons. I highlight two critiques here, and consider a third in Section 4.5.

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: a voter's primary motivation is instrumental, *unless* the instrumental incentives are zero, in which case she votes for her truly favorite candidate.⁴⁶ This modification eliminates any strictly Duvergerian equilibrium. It has no effect, however, on the partially coordinated equilibrium described in Proposition 2.⁴⁷

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 possibility is that all voters perfectly observe a coordinating announcement by a single individual

⁴⁵Anecdotally 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%. Howard's success in the 1997 election was important; at the time of writing he is the leader of Her Majesty's Opposition. Thanks are due to Steve Fisher for this information, and Colin Thompson for his comments on this issue.

⁴⁶Assumption 3 could be extended to specify $U_1 = u_1$ and $U_2 = u_2$ when all others coordinate.

⁴⁷A 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 retort to this response is the argument of Section 3.5: instrumental payoffs may scale up with the size of the electorate, offsetting the fall in the absolute probability of a pivotal outcome.

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.⁴⁸ 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.

4.5. Stability. A further justification for the partially coordinated equilibrium of Proposition 2 is this: an appropriately specified strategy-revision process converges to it. Specifically: beginning from any monotonic strategy with strict multi-candidate support, a sequence of updated best-responses will always lead back to the partially coordinated equilibrium.

To understand this idea, begin with the initial hypothesis that all individuals vote straightforwardly for their preferred candidate. This is equivalent to employing a linear strategy with $a_0 = b_0 = 0$. An individual, acting optimally in response, will employ a linear 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 wider electorate, 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.⁴⁹ Of course, a starting point within the class of monotonic voting strategies with strict multi-candidate support will enter this class within one step (Lemma 4). Formally, I may define an *iterative best-response process* by $b_t = \hat{b}(b_{t-1})$, $a_t = \hat{a}(a_{t-1}, b_{t-1})$. Global stability may be ascertained.⁵⁰ The mapping $\hat{b}(b)$ and associated process $\{b_t\}$ are not contingent on a_t , and hence may be considered in isolation. Lemma 8 in Appendix A confirms that the process $\{b_t\}$ follows a sequence of dampened cycles, and converges to b^* . By inspection, $a_t = \hat{a}(a_{t-1}, b_{t-1})$ converges to $a^* = 0$.

⁴⁸Common 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) The recent literature on global games (Morris and Shin 2003, Myatt, Shin, and Wallace 2002) demonstrates that the lifting of common knowledge assumptions has a dramatic effect on the nature of equilibria. Cox (1997) compared 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. The same is true here.

⁴⁹Earlier work by Fey (1997) considered a process of repeated elections, beginning with an election in which agents act truthfully. I view the iterative best-response process described here as a thought experiment prior to the act of voting, and hence multiple elections are unnecessary.

⁵⁰Of course, the iterative best-response process cannot begin at a fully coordinated voting strategy profile; facing such a strategy, a voter faces no instrumental incentives. Assuming that voters retained the initially postulated strategy profile when this is so would ensure that the process would not move away from a fully Duvergerian equilibrium. On the other hand, the process does not move towards such an equilibrium either; moreover, a fully Duvergerian starting point would, once again, 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 strategy with multi-candidate support, an iterative best-response process converges to the globally stable partially coordinated equilibrium.*

Hence a thought process by which a voter 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 coordination device is necessary: iterative reasoning alone allows an instrumental voter to reach a partially (but not fully) coordinated equilibrium voting strategy.

5. COMPARATIVE STATICS

The strictly Duvergerian equilibria described by Palfrey (1989) and Myerson and Weber (1993) cannot be the subject of comparative-static exercises. Here, however, the partially coordinated equilibrium described in Proposition 2 is amenable to such analysis.

5.1. Strategic Response. A voter’s willingness to vote strategically is captured by b , her response to her private signal of the electoral situation. In the partially coordinated equilibrium of Proposition 2, the solution b^* to the fixed-point equation $b = \hat{b}(b)$ describes this response. As a benchmark, I also consider $\hat{b}(0)$. This represents the response of a voter who acts strategically, but expects all others to vote straightforwardly.⁵¹ To ascertain the response of b^* and $\hat{b}(0)$ to parameter changes, I consider the mapping $\hat{b}(b)$. If $\hat{b}(b)$ is increasing in an exogenous parameter for all b , then (since \hat{b} is decreasing in b) so must be the fixed point b^* as well as $\hat{b}(0) = 2\Phi^{-1}(\gamma)\xi/\kappa^2$.

Examining Equation (2), it is immediate that $\hat{b}(b)$ is increasing in γ : a voter responds more strongly to her signal when greater coordination is required. There is a sense in which γ represents the “safety” of a disliked incumbent office-holder. For instance, when $\gamma \rightarrow 1$ full anti-incumbent coordination is needed. 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, a voter cares only about influencing the outcome, and realizes that, for large γ , coordination behind the leading candidate is crucial. Inspection of Equation (2) also reveals that $\hat{b}(b)$ is decreasing in κ^2 .

Lemma 6. *A voter’s strategic-incentive response to her private signal is increasing in the required qualified-majority γ and the precision of the informative signal $1/\kappa^2$.*

Equation (2) also suggests that $\hat{b}(b)$ is decreasing in the idiosyncrasy of the electorate ξ . This comparative static may be misleading, however. Fixing the proportion $\pi = \Phi(\eta/\xi)$ who prefer the first challenger, the median voter (equivalently, the common component of Assumption 1) η satisfies $\eta = \xi\Phi^{-1}(\pi)$. But this means that, for fixed π , $|\eta|$, and hence the

⁵¹Recall that voting straightforwardly is equivalent to a linear strategy with $a = b = 0$.

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 respond more sluggishly to their private signals; however, this is offset by the increased size of signals received.

5.2. Strategic Incentives. To circumvent this issue, consider the *actual* strategic incentive $\lambda = b\delta$ faced by a voter. The average (both mean and median) incentive is $bE[\delta] = b\eta = \xi\Phi^{-1}(\pi)$. Suppose that candidate $j = 1$ 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 ξ^2 , I must also address the fact that the information precision $1/\kappa^2$ may be influenced by preference idiosyncrasy.

Recall the social-communication micro-foundation of Section 2.4 where a voter observes a sample of m log relative-preferences (including her own) from the electorate. This generated $\kappa^2 = \xi^2/m$, and in this environment m represents the information available to voters. Then

$$\hat{b}(b) = \frac{2\Phi^{-1}(\gamma)\sqrt{m^2 + m(b^2 + 2b)}}{\xi(1+b)} \quad \Rightarrow \quad \hat{b}(0) = \frac{2\Phi^{-1}(\gamma)m}{\xi}. \quad (3)$$

It is easiest to consider the decision-theoretic case; that is, the world in which a strategic voter conjectures that all others will vote straightforwardly. $b = \hat{b}(0)$, and hence the average strategic-incentive is

$$E[\lambda] = \hat{b}(0)\eta = \frac{2\Phi^{-1}(\gamma)m}{\xi} \times \xi\Phi^{-1}(\pi) = 2m\Phi^{-1}(\gamma)\Phi^{-1}(\pi). \quad (4)$$

In this decision-theoretic world the *net* of effect of idiosyncrasy on strategic incentives is zero, and the average strategic incentive depends in a simple way on the required qualified-majority γ , the relative strength of the leading challenger π , and the sample size m .

Turning attention back to b^* , idiosyncrasy continues to have an effect: increased idiosyncrasy tends to increase incentives to vote strategically (Proposition 4, just below). There is an additional factor present. Increased idiosyncrasy means that other voters are (holding their strategic-voting incentives constant) 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.

Proposition 4. *Strategic voting incentives increase with the qualified majority γ , the relative strength of the leading challenger π and the signal precision m . In the uniquely stable equilibrium strategic incentives increase with the idiosyncrasy of the electorate ξ^2 . Also,*

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

The final claim of Proposition 4 reveals the rate at which incentives change when voters are faced with more information. The interesting feature is this: when voters are game-theoretic

(that is, they anticipate strategic voting by others, rather than assuming that others vote straightforwardly) they react *much more slowly* to increased information. In fact, since strategic-voting incentives are increasing (at least asymptotically) with \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.

5.3. The Precision of Information and Duverger’s Law. Allowing $m \rightarrow \infty$ generates a benchmark case where voters are almost perfectly informed. Almost everyone receives a signal close to η , and thus the strategic incentives faced by voters increase without bound as $m \rightarrow \infty$. Hence (for $\pi > 1/2$) almost all voters successfully coordinate on candidate 1. Put simply, this means that the uniquely stable partially coordinated equilibrium is almost perfectly Duvergerian when $m \rightarrow \infty$; a strict version of Duverger’s law (almost) holds when voters have (almost) perfect knowledge of the electoral situation.

Notice that only *one* fully Duvergerian outcome remains as $m \rightarrow \infty$.⁵² Interestingly, the case of $m \rightarrow \infty$ essentially retains the assumption that voters know the electoral situation, but removes the assumption that this is common knowledge. Thus the removal of the common-knowledge element that underpins existing models eliminates all but one equilibrium.

Importantly, the partially coordinated outcome, with multi-candidate support, is *not* related in any way to the non-Duvergerian equilibria of common-knowledge voting games: such equilibria require the *exact* knowledge of underlying candidate-support levels. To see why, notice that if the voting strategy generated $p \neq \frac{1}{2}$ then an infinite strategic-voting incentive would appear as $n \rightarrow \infty$. To avoid this, and hence generate multi-candidate support, p needs to be close (arbitrarily close in a large electorate) to $\frac{1}{2}$. Thus a precisely calculated proportion of the voters who prefer the leading challenger (so for $\pi > \frac{1}{2}$, voters who view candidate 1 as most preferred) must switch *away* from this leader and toward the trailing challenger. Specifically, if a fraction $\pi - \frac{1}{2}$ voters do so, then a tie between the challengers ($p = \frac{1}{2}$) will result in the attenuation of any strategic voting incentives.⁵³ 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 is understood that such non-Duvergerian equilibria require precise common knowledge of the electoral situation: π or equivalently η must be commonly

⁵²The iterative-deletion of weakly dominated strategies employed by Dhillon and Lockwood (2004) does not eliminate equilibria in a common-knowledge version of the qualified-majority voting game.

⁵³In fact, for the case $\pi > \frac{1}{2}$ a slightly greater proportion are required to switch in order to generate a (finite) strategic incentive to switch away from candidate 1.

known by all voters.⁵⁴ These “knife-edge” properties were recognized by Palfrey (1989), who discounted non-Duvergerian equilibria.^{55,56}

5.4. The Impact of Strategic Voting. Changes in ξ^2 influence the strategic incentive faced by voters, at least in the game-theoretic case. They also influence voter preferences, however, and hence potentially determine a voter’s commitment to her most preferred candidate. To assess the impact of strategic voting, I must consider both effects.

To measure this impact, I consider the probability that an individual in an appropriate risk-population votes strategically.⁵⁷ By “risk population” I mean those individuals for whom a strategic vote makes sense. Suppose, that $\pi > 1/2$, or equivalently $\eta > 0$, so that candidate 1 is the leading challenger. A voter who prefers candidate 2 (so that $\tilde{u}_i < 0$) but believes that candidate 1 is the best-placed challenger (so that $\delta_i > 0$, which implies $\lambda = b\delta_i > 0$) faces a positive strategic incentive and is “at risk” of voting strategically.

An easy way to study an at-risk individual is to equip her with a signal satisfying $\delta_i = \eta$ (so that she has an accurate expectation of the electoral situation; of course, she does not know that this the case) and then examine her behavior. She votes for candidate 1 whenever $b\delta_i + \tilde{u}_i \geq 0$, or equivalently (given that $\delta_i = \eta$) when $(1 + b)\eta + \varepsilon_i \geq 0$. Hence the realized support for candidate 1 is⁵⁸

$$p = \Phi\left(\frac{(1 + b)\eta}{\xi}\right) = \Phi((1 + b)\Phi^{-1}(\pi)). \quad (5)$$

Of course, this voter actually prefers candidate 1 with probability π , and so a strategic vote is observed with probability $p - \pi$. Furthermore, she is at risk of voting strategically when she prefers candidate 2 (with probability $1 - \pi$). Hence, the impact of strategic voting,

⁵⁴The exact amount of “wrong-way switching” that is required is sensitive to n ; non-Duvergerian equilibria require common knowledge of the exact electorate size n .

⁵⁵It 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 of this paper. This can be simply regarded as a “fictitious play” thought process in the mind of a player.

⁵⁶Cox (1994, p. 609) noted that non-Duvergerian equilibria “are not unusual in the mathematical sense of being non-generic.” Here, 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.

⁵⁷Using a stripped-down variant of this model, Myatt and Fisher (2002) performed a similar exercise.

⁵⁸Here I have actually given her an artificial signal $\delta_i = \eta$, and so $\varepsilon_i \sim N(0, \xi^2)$. If I had considered her behavior conditional on her actually receiving a signal δ_i , then the distribution of ε_i changes slightly. The reason is that \tilde{u}_i and hence ε_i is a component of the signal δ_i , and so knowledge of δ_i influences the (conditional) distribution of ε_i . Correcting for this has a minimal effect on the analysis (Appendix B).

measured as a proportion of the risk population, is

$$\frac{p - \pi}{1 - \pi} = \frac{\Phi((1 + b)\Phi^{-1}(\pi)) - \pi}{1 - \pi}. \quad (6)$$

Inspection of this expression, coupled with earlier results, yields the following proposition.

Proposition 5. *In the uniquely stable equilibrium, the probability that an at-risk individual votes strategically increases with the required qualified majority, the relative strength of the leading challenger, and the information available to voters; but decreases with idiosyncrasy.*

5.5. Distance-from-Contention and Marginality. A common intuition suggests that strategic voting should be greater in marginal elections, and when the preferred candidate is far from contention. The marginality hypothesis is problematic. The idea behind it is that pivotal events are more likely in more marginal elections.⁵⁹ This idea has no role in a theory of purely instrumental voting, however, since it is the *relative* probability of different pivotal events that matters and not the *absolute* probabilities.

Here, the electoral situation is summarized by γ and π . These parameters are not an ideal way to separate the effects of marginality and distance from contention. Suppose, for instance, that $\gamma > \pi > 1/2$. An increase in π (associated with higher strategic-voting incentives) will increase the gap between the two challengers and hence the distance from contention of candidate 2. Unfortunately this parameter change also closes the gap between candidate 1 and the required qualified-majority, making the election more marginal; it is unclear whether increased strategic-voting is due to marginality or the distance from contention.

A different representation of the electoral situation is required. To do this, I expand consideration to all members of the wider electorate, including those who vote (exogenously) for the disliked $j = 0$. I write the true underlying support for these three candidates as ψ_0 , ψ_1 and ψ_2 where $\psi_0 + \psi_1 + \psi_2 = 1$. This scenario may be easily mapped back into a qualified-majority voting game: the qualified majority of anti-incumbent voters required to defeat $j = 0$ is $\gamma = \psi_0/(1 - \psi_0)$; the true underlying support for candidate 1 among the anti-incumbent electorate is $\pi = \psi_1/(1 - \psi_0)$.

Using this notation, the ideas of marginality and distance from contention may be formalized. Consider an election in which at least some strategic voting is needed to defeat $j = 0$, so that $\psi_0 > \psi_1 > \psi_2$, or equivalently $\gamma > \pi$. Then I may define

$$\begin{array}{l} \text{Winning} \\ \text{Margin} \end{array} = w = \psi_0 - \psi_1 \quad \text{and} \quad \begin{array}{l} \text{Distance from} \\ \text{Contention} \end{array} = d = \psi_1 - \psi_2.$$

⁵⁹Cain (1978, p. 644) provided a classic example of this hypothesis in his analysis of strategic voting in Britain: he expected the pressure to defect to be lower in “noncompetitive” constituencies.

These parameters (together with $\psi_0 + \psi_1 + \psi_2 = 1$) are sufficient to determine the electoral situation. Solving linearly to obtain ψ_0 , ψ_1 and ψ_2 in terms of w and d

$$\left. \begin{array}{l} \psi_0 = (1 + 2w + d)/3 \\ \psi_1 = (1 - w + d)/3 \\ \psi_2 = (1 - w - 2d)/3 \end{array} \right\} \Rightarrow \gamma = \frac{1 + 2w + d}{2 - 2w - d} \quad \text{and} \quad \pi = \frac{1 - w + d}{2 - 2w - d}.$$

Using this formulation, I may change w and d independently, and determine the effect on γ and π . By inspection, both γ and π are increasing in d , and so (as suggested by casual intuition) an increase in the distance from contention results in greater strategic incentives.⁶⁰ The required qualified-majority γ is clearly increasing in the winning margin w . Simple algebra confirms that π is also increasing in w . Taking these two responses to w , I reach the following conclusion: fixing the distance from contention, an increase in the size of the winning margin w (making the election *less* marginal) actually *increases* the incentives to vote strategically.

Proposition 6. *When coordination is required to defeat $j = 0$ ($\psi_0 > \psi_1 > \psi_2$) the incentive to vote strategically increases with both the winning margin and the distance from contention.*

This prediction runs against established intuition. After controlling for the distance from contention, strategic voting should be *lower* in more marginal elections.⁶¹

5.6. The Importance of Strategic Voting and Election Inversion. Equation (6) measures the *impact* of strategic voting. In a similar way, I may assess the *importance* of strategic voting. An inversion of Equation (5) would yield π as a function of p : $\pi = \Phi(\Phi^{-1}(p)/(1+b))$. This reveals what the election outcome *would* have been in the *absence* of strategic voting. If $p > \gamma > \pi$, then strategic voting changes the identity of the winning candidate. For a more accurate assessment of this effect, I must also account for the variation in voter's signals.

⁶⁰Empirically the distance from contention is well known as a strong predictor of strategic voting, allowing the measure to become a basis to construct validity tests for different measures of strategic voting (Franklin, Niemi and Whitten 1992, 1993, 1994 and Evans and Heath 1993, 1994).

⁶¹To adjudicate between the common intuition of Cain (1978) and others, and the comparative static results presented here, empirical analysis is required. In recent work, Fisher (2000) found that strategic voting increases with the winning margin in UK General Elections. The effect is weak, but nevertheless is significant at the 10% level when estimated across the three elections of 1987, 1992 and 1997. This analysis was extended by Myatt and Fisher (2003) and Fisher (2001b). In the former paper, a version of the strategic incentive λ was shown to explain the pattern of strategic voting across all three elections, and its inclusion eliminated the significance of the winning margin and distance from contention as separate measures. In the latter paper, Fisher (2001b) included a range of other explanatory variables, including political interest, education, party identification, and local campaign-spending. A version of λ continued to have strong explanatory power. This means that there is strong empirical support for the present theory relative to informal intuition. Interestingly, this provides a partial response to the critique of rational-choice theory offered by Green and Shapiro (1994): the formal theory offers explanations that differ from intuition, and yet better fit the data.

Straightforward manipulations lead to

$$p = \Phi \left(\frac{(1+b)\eta}{\sqrt{\text{var}[b\delta_i + \tilde{u}_i | \eta]}} \right) = \Phi \left((1+b)\Phi^{-1}(\pi) \sqrt{\frac{m}{(m-1) + (1+b)^2}} \right)$$

$$\Rightarrow \pi = \Phi \left(\frac{\Phi^{-1}(p)}{1+b} \sqrt{\frac{(m-1) + (1+b)^2}{m}} \right). \quad (7)$$

Equation (7) allows me to “invert” an election outcome (that is, I remove the effect of strategic voting) and requires as input the parameters γ , p and m . The parameters γ and p summarize the realized electoral situation.⁶² In the calibration exercise of Section 5.7, I examine the response of strategic voting to the information parameter m . Before doing this, however, I must consider the specification of ξ^2 .

ξ^2 measures the idiosyncrasy of preferences across the electorate. An increase in ξ^2 results in a larger number of voters who are more heavily committed to their preferred candidate (for instance, when u_{i1}/u_{i2} is large). As I show in Appendix B, a value for ξ^2 may be pinned down by the preferences of the median supporter of a particular challenging candidate.

To understand this idea, consider the median supporter of candidate 1, and suppose that she prefers candidate 1 twice as much as candidate 2, so that $u_1 = 2u_2$. Since she is the median supporter, it follows that $\Pr[u_1 \geq 2u_2] = \pi/2$. Together with π , this equation is sufficient solve for ξ^2 . In fact, for this example and $\pi = 1/2$, it solves for $\xi^2 = 1.056$ (see the Appendix). I thus fix ξ^2 at this value for the calibration exercises of Sections 5.7 and 6.3.

5.7. Application: New York (1970). I am now in a position to calibrate the model and examine the impact and importance of strategic voting in the context of reasonable parameter values. I return to the case of the 1970 New York senatorial election (Table 1). Mapping this example into the model yields the parameters γ and p ,

$$\gamma = \frac{\psi_0}{\psi_1 + \psi_2} = \frac{2,288,190}{3,605,704} = 0.635 \quad \text{and} \quad p = \frac{\psi_1}{\psi_1 + \psi_2} = \frac{2,171,232}{3,605,704} = 0.602.$$

Hence liberal voters needed to achieve a qualified majority of $\gamma = 0.635$ to defeat the disliked Buckley and yet achieved only $p = 0.602$; this split in the liberal vote allowed Buckley to win. Following the analysis above, I fix the idiosyncrasy parameter at $\xi^2 = 1.056$. For any specified m , I may now calculate the equilibrium response b of a voter to her signal δ_i . For a variety of values for π , I calculate the impact of strategic voting, by substituting into Equation (6). The results are displayed in Figure 2(a). As expected, the impact of strategic voting on the risk population increases with both m and π . For small m and π , strategic voting is limited. However, it takes only moderate values of m to dramatically increase the

⁶²This contrasts with the true underlying electoral situation, which relies on π rather than p .

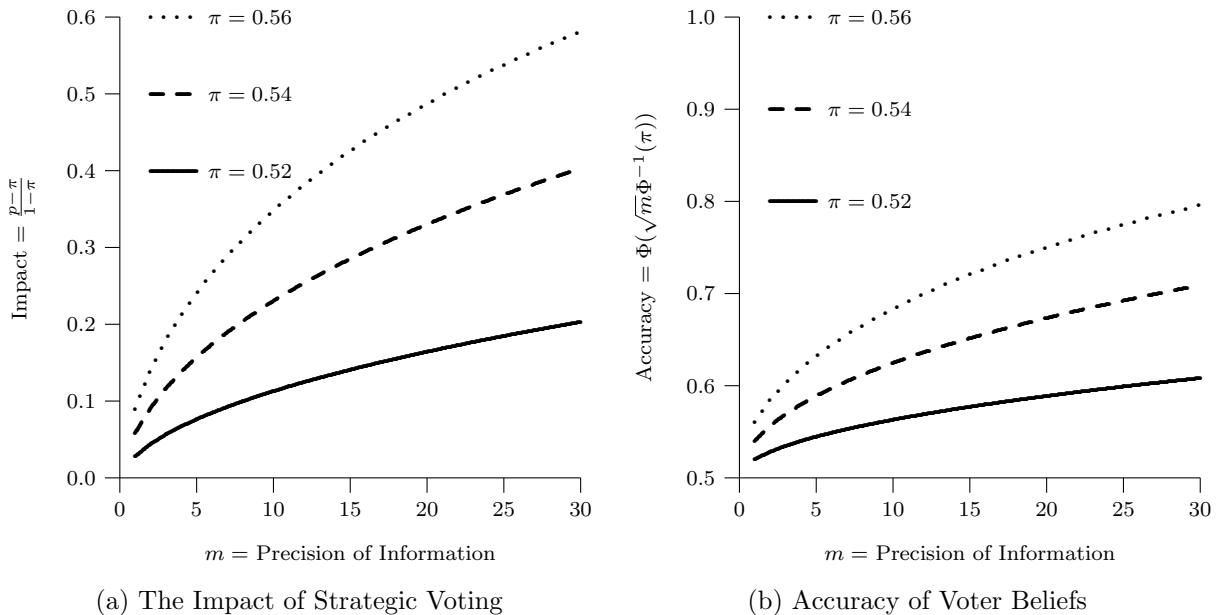


FIGURE 2. Strategic Voting in the New York Senatorial Election

equilibrium impact of strategic voting to (when π is relatively large) very high levels. Are such large values for m reasonable?

Adopting the social-communication micro-foundation for the informative signal δ_i described in Section 2.4, it would seem that moderately large values for m are appropriate: an individual voter might be expected to interact with a large number of people during day-to-day life. This may be misleading, however. Even if the number of individuals observed is large, the *effective* value for m will be rather lower. I offer two reasons here. First, the observation of the preferences of others is likely to occur with noise. Second, and most importantly, if a voter interacts with individuals whose preferences are correlated, then the observation of additional voters is likely to add little extra information; I address this second issue formally in Section 6. To cope with these issues here, I may check the plausibility of m by computing the *accuracy* of beliefs.

As in Section 2.5 I define the accuracy of beliefs as the probability that the informative signal correctly identifies the leading challenger. I continue to assume (without loss of generality) that $\eta > 0$, so that $\pi > \frac{1}{2}$. A signal indicates the correct leader when $\delta_i > 0$. Recall that this occurs with probability $\alpha = \Pr[\delta_i \geq 0] = \Phi((\sqrt{m}/\xi)\eta) = \Phi(\sqrt{m}\Phi^{-1}(\pi))$. This last formula allows me to cross-check m against the accuracy of beliefs. For instance,

$$m = 25 \quad \Rightarrow \quad \alpha = \Phi(\sqrt{m}\Phi^{-1}(\pi)) = \begin{cases} 0.599 & \text{if } \pi = 0.52, \\ 0.692 & \text{if } \pi = 0.54, \text{ and} \\ 0.775 & \text{if } \pi = 0.56, \end{cases}$$

with further values illustrated in Figure 2(b). This calibration exercise demonstrates that if the outcome of the New York senatorial election is to be consistent with the present model then either the true value for π must be relatively close to $\frac{1}{2}$, or the information available to voters must not identify the challenger with very high probability.

In fact, I may calculate the parameter π that is consistent with the data by using Equation (7). Doing so with $m = 25$ yields $\pi = 0.528$. This suggests that, absent strategic voting, Ottinger’s true support may have been closer to $0.528 \times 61\% \approx 32.2\%$. Of course, the value of $m = 25$ implies that voters’ beliefs were not particularly accurate, and hence this “inversion” procedure is subject to the critique that an instrumental theory requires voters to have relatively “fuzzy” beliefs. I turn attention to this issue.

6. COMMUNITY SAMPLING AND FUZZY BELIEFS

6.1. Social Communication within Communities. The calibration exercise for the New York election suggests that the present theory is consistent with observed data only when the information available to voters is limited. If I were to introduce the possibility of opinion polls, then the probability that a voter will be able to identify the lead challenger will be relatively large. In such situations, therefore, the theory may well predict too much strategic voting, leading me to question the assumption of instrumental motivations.

Turning to other scenarios, however, this is not necessarily the case. The plurality rule is used in the parliamentary constituencies of the United Kingdom. At a constituency level, opinion polls are not typically available (see Footnote 31), and hence a voter must learn of the electoral situation via social communication. If I am to take seriously the sampling procedure described in Section 2.4, then I must also recognize that voters are likely to communicate with other individuals who are similar to them.

I formalize this idea with a notion of *community communication*. Suppose that each individual belongs to a community, and that a fraction ω of her idiosyncratic component is shared with all individuals in the community. This is modelled via the following decomposition:

$$\varepsilon_i = \theta + \tilde{\varepsilon}_i \quad \Rightarrow \quad \tilde{u}_i = \eta + \theta + \tilde{\varepsilon}_i \quad \text{where} \quad \begin{aligned} \text{var}[\theta | \eta] &= \omega\xi^2, & \text{and} \\ \text{var}[\tilde{\varepsilon}_i | \eta] &= (1 - \omega)\xi^2. \end{aligned}$$

Thus preferences are equal to a common component across the electorate, plus a community component and then a further individual component; a multi-level model.⁶³ The parameter ω represents the relative importance of the community effect. This specification may seem reasonable; after all, the preferences of the electorate may well be affected by events at a community level. Now, suppose that a voter is only able to sample the preferences of

⁶³A related specification for a model of district-based electoral systems was proposed by Gelman and King (1994); such a hierarchical probit structure was also used by Gelman, Katz, and Tuerlinckx (2002, pp. 431-33) in their study of power indices.

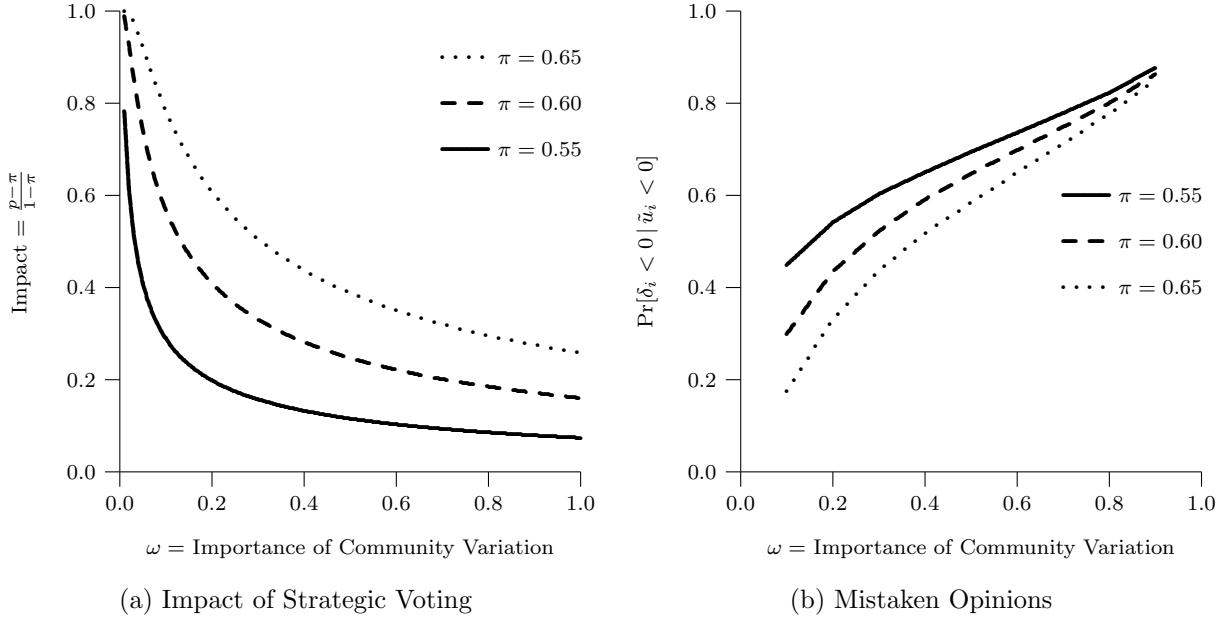


FIGURE 3. The Effect of Community Variation

individuals from her own community. Since θ is common to every sample member, it cannot be averaged out. This means that $\text{var}[\delta_i | \eta] \geq \text{var}[\theta | \eta]$. In fact, the best that a voter can do is to identify perfectly the typical member of her community. This would be equivalent to the observation of $\eta + \theta$, and hence a signal $\delta_i = \eta + \theta$. In this case,

$$\text{var}[\delta_i | \eta] = \text{var}[\theta | \eta] = \omega \xi^2 \quad \Rightarrow \quad m = \frac{\text{var}[\delta_i | \eta]}{\xi^2} = \frac{1}{\omega}.$$

It follows that there is an upper bound to the precision of information. For instance, if $\omega = 0.2$ so that 20% of individual variation is due to variation across communities, then $m \leq 5$. This suggests that appropriate values for m might well be very low indeed. In Figure 3(a) I plot the impact of strategic voting against ω . By inspection, even a small element of community variation dramatically reduces the impact of strategic voting.

6.2. The False-Consensus Effect. The presence of community variation has additional implications for the pattern of correlation between a voter's preferences and her beliefs. Suppose that candidate 1 is the true leading challenger ($\eta > 0$) but that a particular voter i prefers candidate 2, so that $\tilde{u}_i < 0$. When ω is large, so that the individual variation is small relative to community variation, then it is highly likely that voter i is drawn from a community where $\eta + \theta < 0$. If her opinions are based on the observation of her community (so that $\delta_i = \eta + \theta$ as above) then she will believe that her favored candidate 2 is the leading challenger. In other words, for large ω , voters are likely to live in communities where their opinions are shared, and hence any information on the electoral situation will be highly

correlated with their own preferences. To express this formally, I may examine

$$\Pr[\delta_i < 0 \mid \tilde{u}_i < 0] = \frac{1}{1 - \pi} \int_{\Phi^{-1}(\pi)/\sqrt{\omega}}^{\infty} \Phi\left(\frac{z\sqrt{\omega} - \Phi^{-1}(\pi)}{\sqrt{1 - \omega}}\right) d\Phi(z). \quad (8)$$

This is the probability that a voter’s signal indicates that candidate 2 is the lead challenger given that she prefers candidate 2. The equality is derived in Appendix B.

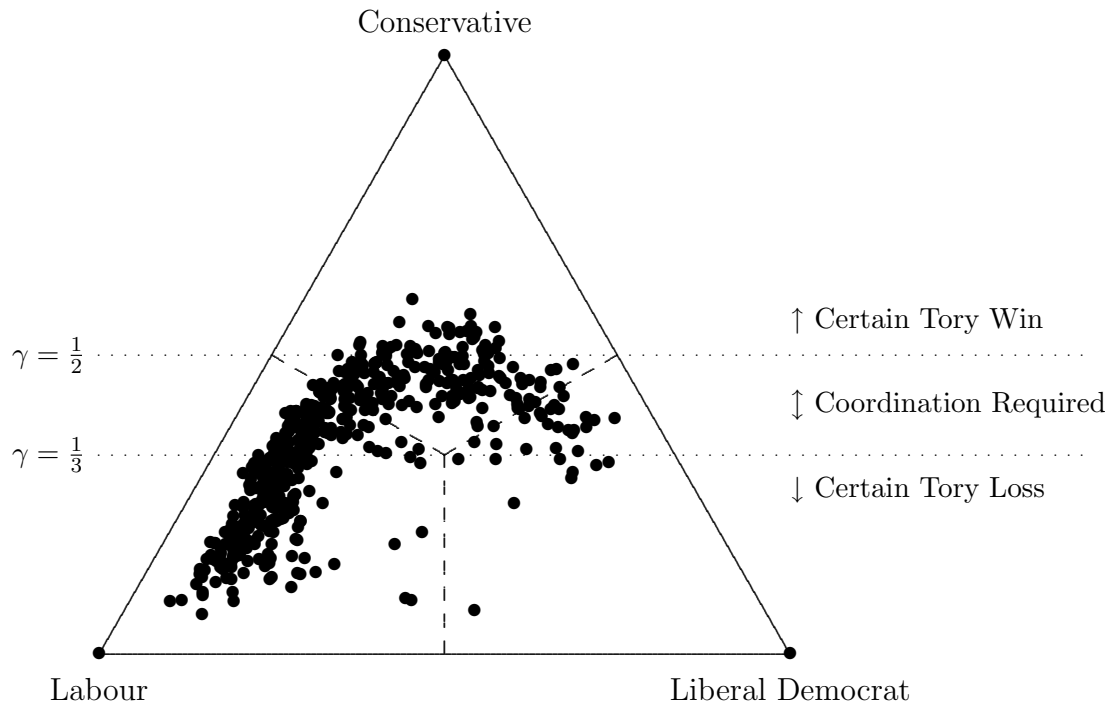
I plot this probability against ω in Figure 3(b). Since I have, without loss of generality, specified $\pi > 1/2 \Leftrightarrow \eta > 0$, this is equivalent to the probability that a supporter of the trailing challenger is mistaken in her belief of this candidate’s likely prominence. By inspection, for large ω this probability is quite high, and importantly it exceeds $\frac{1}{2}$.

I make two observations. First, this analysis shows that an exaggerated belief (referred to as the “false-consensus effect” in the psychology literature) in the likely success of a preferred candidate is in no way irrational.⁶⁴ In fact, voters form their opinions in a way that is entirely rational; any “exaggerated belief” is merely a consequence of the fact that preferences and information are highly correlated. Second, it suggests that if community variation is large then one might expect to see a large incidence of apparently (but not actually) misplaced confidence in preferred candidates.

6.3. Application: United Kingdom (1997). The fully Duvergerian equilibria of a “common knowledge” plurality-voting game predict the strict bipartism at the level of election district; all voters coordinate behind two leading candidates. The 1970 New York election provides a counter-example to both strict bipartism and the non-Duvergerian equilibria emphasized by Cox (1994). The data’s rejection of such common-knowledge equilibria is more widespread; I return to the UK General Election of 1997 for a further illustration. In 1997, the three major political parties (Conservative, Labour and Liberal Democrat) competed throughout 527 English parliamentary constituencies. In Figure 4, I use barycentric coordinates to plot their relative vote shares: both strictly Duvergerian and non-Duvergerian equilibria are absent.

Using the idea of community communication, I now consider calibrations of the model with rather less information, corresponding to a lower values of m . Naturally, lower values of m lead to lower accuracy, and hence voters find it harder to identify the true leading challenger to the disliked incumbent. To take an example, suppose that $\omega = 1/5$, or equivalently $m = 5$.

⁶⁴Goeree and Grosser (2003) attributed the “false consensus” terminology to Ross, Greene, and House (1977), and credited Dawes (1990) with the observation that is not justified automatically. They noted that “[w]hen someone’s decision is driven by their taste, which is seen as a draw from a population, then it is perfectly rational to use this decision the same way as any other random sample of size one. Only when information about one’s own taste is outweighed [sic] is the perceived consensus false.” Notice that the false-consensus effect differs from the confirmatory-bias modelled by Rabin and Schrag (1999); they considered over-confidence in hypotheses rather than a correlation between an individual’s preferences and her beliefs.



The relative vote shares for the three major parties for 527 English constituencies are plotted using a *barycentric* or *simplex* plot. The three corners of the simplex represent 100% vote shares for the labelled party. A bullet point “•” indicates a constituency. Its location on the graph is a weighted average of the three extreme points, with weights corresponding to the relative vote shares of the major parties. The sides of the simplex (solid lines) represent points where only two parties receive votes. The hatched lines represent points where two parties tie for the lead, and hence separate the “win zones” for each party. The dotted lines delineate the constituencies in which the Conservative party polled between 1/3 and 1/2 of the votes for major parties. These are the 270 constituencies in which coordination of anti-Conservative voters is required to avoid a Tory win.

FIGURE 4. British General Election 1997

If $\pi = 0.635$, then a voter correctly identifies the leading challenger with probability 78%, and hence 22% of the electorate are mistaken in their perception of the leading challenger.

Does this resonate with empirical observation? To assess this, I may turn to the 1997 British Election Survey. Fisher (2000, p. 6) commented that 68.5% of those voters who expected their preferred party to come second actually found that it came third.⁶⁵ This suggests that British voters may have rather inaccurate opinions of their constituency’s characteristics, and hence the electoral situation. More interestingly, he found that roughly half of those

⁶⁵This was an empirical instance of the false-consensus effect. Brown (1982) and Baker, Koestner, Warren, Losier, and Vallerand (1995) reported similar findings in US and Canadian elections.

	Actual	$\omega = 0.05$	$\omega = 0.1$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$
Lab > Con > Lib	118	57	70	82	105	105	105
Con > Lab > Lib	73	134	121	109	86	86	86
Con > Lib > Lab	54	70	68	65	65	63	59
Lib > Con > Lab	25	9	11	14	14	16	20
Con \rightarrow Lab	—	61	48	36	13	13	13
Con \rightarrow Lib	—	16	14	11	11	9	5
Con Seats Lost	—	77	62	47	24	22	18
Impact	—	35.8%	31.9%	26.3%	21.9%	18.1%	14.6%

The first four rows of the table classify the 270 seats according to the relative vote shares of the three main parties. I have excluded constituencies in which the Conservative party came third, hence there are four possible configurations. The first column gives the classification according to the actual results. The remaining columns give the results when the strategic voting element is removed, for values of ω from 0.05 to 0.5 (and hence m from 2 to 20). The fifth, sixth, and seventh rows give the number of seats lost by the Conservatives due to strategic switching. The final row gives the probability of a strategic vote by a typical member of the risk-population; that is, someone who prefers the trailing challenger and is equipped with a correct signal realization indicating the electoral situation (Equation (6)).

TABLE 2. Inverted Election Results: English Constituencies 1997

whose preferred party came third expected it to come second. This coincides rather closely with the analysis of Section 6.2. It would appear, therefore, that a situation in which voters' beliefs are somewhat "fuzzy," due to significant community-specific effects in preferences, generates (at first blush) the kind of results observed in survey data.

Employing these ideas, I turn to the data displayed in Figure 4. As a calibration exercise, I consider the 270 constituencies where the Conservatives polled between 1/3 and 1/2 of the vote, and identify them as unpopular incumbents.⁶⁶

For each constituency, I calculate appropriate values for γ and p . For a variety of values of ω (and hence m) I calculate b and, using Equation (7), "invert" the election result to obtain the notional true level of support for each challenging candidate. Using these notional levels of support, I then calculate the election results in the absence of strategic voting.

Of course, this procedure is very much a "back-of-the-envelope" calibration exercise, and should be seen as nothing more than indicative. Nevertheless, the results are interesting,

⁶⁶The implicit assumption is that all supporters of the Labour and Liberal Democrat parties rank the Conservative party last. Of course, this is not actually the case, and in fact a significant proportion of Liberal Democrat voters ranked the Labour party last in 1997. Nevertheless, this simplification is used to generate a "ball-park" figure for the amount of strategic voting that is consistent with observed patterns.

and given in Table 2. For $\omega = 0.2$, so that community variation accounts for 20% of the idiosyncratic variation of voter preferences, the results suggest that the Conservative party lost 47 seats due to strategic voting towards the Labour and Liberal Democrat parties. Furthermore, this same parameter choice suggests that, according to the theory, strategic voting affected around 26% of the “risk population” of trailing candidate supporters. This matches well with the empirical estimates of Fisher (2000). He estimates (again using BES data) that 24.4% of votes in the risk population of third-party supporters voted strategically in the 1997 British General Election.⁶⁷ The overall incidence of strategic voting from a calibrated model appears to be about right.

7. CONCLUDING REMARKS

I began with the observation that many analyses of plurality-rule elections predict the complete coordination of strategic voting, and hence support for only two candidates: a strictly bipartite outcome. My suggestion that stable multi-candidate support will arise as part of a fulfilled-expectations equilibrium follows from the weakening of a single assumption: I have removed the common knowledge of the electoral situation. In doing so, I have attempted to follow a suggestion of McKelvey and Ordeshook (1985, p. 57) by investigating “the extent to which the equilibria of systems with limited information corresponds to the equilibria of systems with full information.” This investigation yields a uniquely stable partially coordinated voting equilibrium in which there is some, but not complete, strategic vote-switching toward a leading challenger. In the limit, as voters’ information becomes precise, this equilibrium converges to one of the strictly Duvergerian equilibria described by Palfrey (1989) and Myerson and Weber (1993); nevertheless, the community-communication story suggests that, in some settings, there will be a limit to the precision of information available to voters, and hence a bound to Duverger’s psychological effect.

Formally, the equilibrium described here is radically different from the non-Duvergerian equilibria that were important to Cox (1994). Despite this, the uniquely stable partially coordinated equilibrium incorporates *precisely* the informal (and insightful) intuition that he offered. For instance, Cox (1997, p. 86) interpreted the 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 required that “it was not clear who was in third and who in second” and that “neither [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 a *precisely calculated* strategic swing. This intuition, however,

⁶⁷This impact figure was calculated across all English constituencies and not just those considered here.

is entirely consistent with the partially coordinated equilibrium described here: voters are uncertain of the ranking of challenging candidates.

I conclude by noting that this paper offers insights and predictions that differ somewhat from casual intuition; for instance, I predict the negative-feedback effect, and the rejection of the marginality hypothesis described by Cain (1978) and others. While I have made no attempt to conduct an empirical study, some of the predictions do seem consistent with the outcomes of UK elections; moreover, related work takes some of the predictions to the data, and they are not rejected. This is in contrast to the claim of Green and Shapiro (1994, p. 6) that rational-choice theory does “little more than restate existing knowledge in rational-choice terminology.” It may be argued that results offered here yield a response to this critique.

APPENDIX A. OMITTED PROOFS

Proof of Lemma 1. By Assumption 2, and conditional on η , the sufficient-statistic signal δ_i and the log-relative preference are joint normal. Consider the construction of an alternative signal $\tilde{\delta}_i = \alpha\delta_i + (1 - \alpha)\tilde{u}_i$ for some $\alpha \in \mathbb{R}$. This satisfies $E[\tilde{\delta}_i | \eta] = \eta$, and has variance $\text{var}[\tilde{\delta}_i | \eta] = \alpha^2\kappa^2 + (1 - \alpha)^2\xi^2 + 2\alpha(1 - \alpha)\text{cov}[\tilde{\delta}_i, \tilde{u}_i | \eta]$. Choosing α to minimize this,

$$\frac{\partial \text{var}[\tilde{\delta}_i | \eta]}{\partial \alpha} = 2\alpha\kappa^2 - 2(1 - \alpha)\xi^2 + 2(1 - 2\alpha)\text{cov}[\tilde{\delta}_i, \tilde{u}_i | \eta] = 0.$$

If δ_i is a sufficient statistic, then this must hold at $\alpha = 1$. This yields $\text{cov}[\tilde{\delta}_i, \tilde{u}_i | \eta] = \kappa^2$. \square

Proof of Lemma 2. Follows from the argument presented in the main text. \square

Proof of Lemma 3. I need to show that p takes on all values in $(0, 1)$ interval with positive density. First, $v(\delta_i, \tilde{u}_i) = 1$ for some pair δ_i and \tilde{u}_i , by assumption. By monotonicity, it retains this value for larger δ_i and \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 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

Sketch Proof of Proposition 1. Examine Equation (1) and notice that 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. For a rigorous proof of this proposition, see the not-for-publication Appendix C. \square

Proof of Lemma 4. Suppose that all voters use a monotonic voting strategy exhibiting strict multi-candidate support. Define $p = \mathbb{E}[v(\delta_i, \tilde{u}_i) | \eta] = H(\eta)$. This is strictly and smoothly increasing in η . Write $h(\eta) = H'(\eta)$. Following the receipt of a signal δ ,

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

where $\Phi(\cdot)$ is the cumulative distribution function of the normal. This uses the fact that private posterior beliefs over η are normal. Differentiating,

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

Turning to the strategic incentive λ ,

$$\begin{aligned} \log \frac{f(\gamma | \delta)}{f(1 - \gamma | \delta)} &= \log \frac{h(H^{-1}(1 - \gamma))}{h(H^{-1}(\gamma))} - \frac{(H^{-1}(\gamma) - \mathbb{E}[\eta | \delta])^2}{2\text{var}[\eta | \delta]} + \frac{(H^{-1}(1 - \gamma) - \mathbb{E}[\eta | \delta])^2}{2\text{var}[\eta | \delta]} \\ &= \log \frac{h(H^{-1}(1 - \gamma))}{h(H^{-1}(\gamma))} + \frac{(H^{-1}(1 - \gamma))^2 - (H^{-1}(\gamma))^2}{2\text{var}[\eta | \delta]} + \frac{H^{-1}(\gamma) - H^{-1}(1 - \gamma)}{\text{var}[\eta | \delta]} \mathbb{E}[\eta | \delta]. \end{aligned} \quad (9)$$

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

Proof of Lemma 5. Under linear strategies, individual i votes for candidate 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 $\Phi(\cdot)$ 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 (9). 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 (9),

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

Similarly,

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

and so 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

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

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

$$\begin{aligned} \log \frac{f(\gamma|\delta)}{f(1-\gamma|\delta)} &= \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)(a + (1+b)\delta)}{\kappa^2(1+b)^2} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)a}{\kappa^2(1+b)^2} + \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)\delta}{\kappa^2(1+b)} \\ &= \frac{a\hat{b}(b)}{(1+b)} + b\delta \quad \text{where} \quad \hat{b}(b) = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma)}{\kappa^2(1+b)}. \end{aligned}$$

This yields the desired intercept parameter. For the slope parameter,

$$\tilde{k}^2 = \text{var}[\varepsilon_i + b(\delta_i - \eta) | \eta] = \xi^2 + b^2\kappa^2 + 2b\text{cov}[\delta_i, \tilde{u}_i | \eta] = \xi^2 + (b^2 + 2b)\kappa^2,$$

where the last step follows from Lemma 1. Substituting, this yields

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

For the final part of the proof, differentiate this with respect to b :

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

For $b \geq 0$, this derivative is negative if and only if

$$1 \geq \frac{(1+b)^2\kappa^2}{\xi^2 + (b^2 + 2b)\kappa^2} = \frac{(b^2 + 2b + 1)\kappa^2}{\xi^2 + (b^2 + 2b)\kappa^2} \quad \Leftrightarrow \quad \kappa^2 \leq \xi^2,$$

which holds since ξ^2 is an upper bound to the variance of a voter's signal. Finally, evaluate $\hat{b}(b)$ at $b = 0$ and as $b \rightarrow \infty$ to obtain the final statements of the lemma. \square

Proof of Lemma 6. By inspection of Equation (2). \square

The following lemma yields the proof of Proposition 2.

Lemma 7. *The mapping $\hat{b}(b)$ has a unique fixed point $b^* > 0$. This is bounded as follows*

$$2\Phi^{-1}(\gamma) \leq \kappa b^* \leq \Phi^{-1}(\gamma) \sqrt{2 + 2\sqrt{\frac{(\Phi^{-1}(\gamma))^2 + \xi^2}{(\Phi^{-1}(\gamma))^2}}}, \quad (10)$$

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

Proof of Lemma 7. The uniqueness of b^* follows from the fact that $\hat{b}(b)$ is decreasing. To obtain an upper bound for the fixed point b^* , write $\hat{b}(b)$ as

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

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

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

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 the desired upper bound from this, 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

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

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

Proof of Proposition 2. From Lemma 7 and the argument given in the main text. \square

The following lemma yields the proof of Proposition 3.

Lemma 8. *b^* is globally stable in the iterative best response dynamic: $b_t \rightarrow b^*$ as $t \rightarrow \infty$.*

Proof of Lemma 8. 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. This satisfies

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

Both 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 3. 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 8 completes the proof. \square

Proof of Proposition 4. The effects of γ and m follow immediately from Lemma 6. 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 (3) 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 8). It remains to consider the behavior of strategic incentives as $m \rightarrow \infty$. The first equality is a restatement of Equation (4). 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 7 demonstrates that κb^* attains the upper bound of Equation (10) as $\kappa \rightarrow 0$. This completes the proof. \square

Proof of Proposition 5. From the discussion in the text and properties of b^* . \square

Proof of Proposition 6. By inspection γ is increasing in both w and d and π is increasing in d . Differentiation of π with respect to w demonstrates that π is also increasing in w . \square

APPENDIX B. OMITTED MATERIAL FROM SECTIONS 5– 6.

B.1. Refining the Impact Measures. When calculating the impact of strategic voting (via Equation (6)) I “equipped” a voter with an accurate signal realization $\delta_i = \eta$, and then calculated $p = \Phi((1 + b)\Phi^{-1}(\pi))$. A slightly different result is obtained when I consider her behavior conditional on her actually receiving a signal realization $\delta_i = \eta$. This is because her preferences form part of the signal, and hence, conditional on $\delta_i = \eta$, it is no longer the case that $\tilde{u}_i \sim N(\eta, \xi^2)$. In fact, Bayesian updating following observation of $\delta_i = \eta$ yields

$$\tilde{u}_i | \eta, \delta_i = \eta \sim N\left(\eta, \frac{\xi^2(m-1)}{m}\right) \Rightarrow p = \Phi\left(\sqrt{\frac{m}{m-1}} \frac{(1+b)\eta}{\xi}\right),$$

which is a slight modification to Equation 6.

B.2. Specification of the Idiosyncrasy Parameter ξ^2 . ξ^2 may be specified in a number of different ways. One way is to set two different quantiles for voters with different intensity of preference. For instance, a voter prefers candidate 1 when $u_1 \geq u_2$, and hence I may define $\pi(1) = \Pr[u_1 \geq u_2] = \Phi(\eta/\xi)$. A voter prefers candidate 1 k as much as candidate 2 (relative to the disliked status-quo) when $u_1 \geq ku_2 \Leftrightarrow \tilde{u} \geq \log k$, yielding $\pi(k) = \Pr[u_1 \geq ku_2] = \Phi((\eta - \log k)/\xi)$. These two equations are sufficient to tie down η and ξ^2 :

$$\begin{aligned} \pi(k) = \Phi((\eta - \log k)/\xi) &\Rightarrow \xi\Phi^{-1}(\pi(k)) = \eta - \log k = \xi\Phi^{-1}(\pi(1)) - \log k \\ &\Rightarrow \xi = \frac{\log k}{\Phi^{-1}(\pi(1)) - \Phi^{-1}(\pi(k))}. \end{aligned}$$

To see this formula in action, consider an electorate where a fraction $\pi = \pi(1) = 0.6$ of the electorate rank candidate 1 highest, and half of these (or a fraction $\pi(2) = 0.3$ of the electorate) prefer candidate 1 twice as much as candidate 2. Then,

$$\xi = \frac{\log 2}{\Phi^{-1}(0.6) - \Phi^{-1}(0.3)} = 0.891 \Rightarrow \xi^2 = 0.794.$$

Using this formula, ξ^2 may also be specified by considering the median supporter of candidate 1. If the median supporter prefers candidate 1 k times as much as candidate 2, then (by definition of being the median in this group) $\pi(k) = \pi(1)/2$. Hence $\xi = \log k / [\Phi^{-1}(\pi) - \Phi^{-1}(\pi/2)]$. When $\pi = 1/2$ this formula reduces to $\xi = -\log k / \Phi^{-1}(1/4)$. Illustrating this last formulation, consider a balanced electorate in which $\pi = 1/2$. Suppose that the median supporter of candidate 1 prefers her favored candidate twice as much as candidate 2. Then $\xi = 1.028$ and $\xi^2 = 1.056$. For $k = 1.5$ and $k = 2.5$ the outcomes are $\xi = 0.6$ and $\xi = 1.36$ respectively. Hence a range of $0.5 < \xi < 1.5$ might seem appropriate for the idiosyncrasy parameter.

B.3. Calibration of the UK General Election of 1997. I considered only the 529 English constituencies. The Speaker’s seat and Tatton (where the major parties pulled out as part of an anti-sleaze protest) are excluded, yielding 527 remaining constituencies. The

three leading candidates were then the three major political parties. Writing ψ_0 for the Conservative vote share among these three, I excluded $\psi_0 \geq 1/2$ (a certain Tory win) and $\psi_0 \leq 1/3$ (a certain Tory loss). Within the remaining 270 constituencies I labelled the leading challenger as candidate 1 and the trailing challenger as candidate 2. This led to $\gamma = \psi_0/(1 - \psi_0)$ and $p = \psi_1/(\psi_1 + \psi_2)$. Using $m = 1/\omega$, I calculated the response parameter b . I then employed Equation (7).

APPENDIX C. NOT-FOR-PUBLICATION

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

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

where $f(p)$ is a continuous density taking strictly positive values on $(0, 1)$, following Lemma 3. Dependence on δ is suppressed for simplicity. I next introduce the notation $G(p)$,

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

$G(p)$ 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 I fix $\min\{1/4, 1 - \gamma\} > \epsilon > 0$. For convenience, I now define $f_\epsilon(x) = \max_{x-\epsilon \leq p \leq x+\epsilon} f(p) < \infty$, where this is well defined since $[x - \epsilon, x + \epsilon]$ is a compact set and $f(p)$ is continuous from Lemma 3. I formulate an upper bound for the ratio $r(\tilde{\gamma})$,

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

The right hand side of this equation has five terms. I consider them in turn. First,

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

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

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

so that taking the second term, and allowing $n \rightarrow \infty$, it follows that

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

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

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

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

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

A similar procedure bounds the limit below, and hence $r(\tilde{\gamma}) \rightarrow f(\tilde{\gamma})$. Next, I construct a compact interval $[\gamma - \eta, \gamma + \eta]$ around γ , for small η . For $\tilde{\gamma} \in [\gamma - \eta, \gamma + \eta]$, 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 $\gamma_n \in [\gamma - \eta, \gamma + \eta]$ for sufficiently large n . For sufficiently large n , $r(\gamma_n)$ is arbitrarily close to $f(\gamma_n)$. By taking η 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

REFERENCES

- ALDRICH, J. H. (1993): "Rational Choice and Turnout," *American Journal of Political Science*, 37(1), 246–78.
- BAKER, L., R. KOESTNER, N. M. WORREN, G. F. LOSIER, AND R. J. VALLERAND (1995): "False Consensus Effects for the 1992 Canadian Referendum," *Canadian Journal of Behavioural Science*, 27(2).
- BECK, N. (1975): "A Note on the Probability of a Tied Election," *Public Choice*, 23, 75–9.
- BLACK, D. (1948): "The Decisions of a Committee using a Special Majority," *Econometrica*, 16(3), 245–61.
- (1958): *The Theory of Committees and Elections*. Cambridge University Press, Cambridge, UK.
- BROWN, C. E. (1982): "A False Consensus Bias in 1980 Presidential Preferences," *Journal of Social Psychology*, 118, 137–38.

- CAIN, B. E. (1978): "Strategic Voting in Britain," *American Journal of Political Science*, 22(3), 639–655.
- CAPLIN, A., AND B. NALEBUFF (1988): "On 64%-Majority Rule," *Econometrica*, 56(4), 787–814.
- CARLSSON, H., AND E. VAN DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.
- CHAMBERLAIN, G., AND M. ROTHSCHILD (1981): "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory*, 25, 152–62.
- COX, G. W. (1984): "Strategic Electoral Choice in Multi-Member Districts: Approval Voting in Practice?," *American Journal of Political Science*, 28(4), 722–38.
- (1987): "Duverger's Law and Strategic Voting," University of California, San Diego.
- (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.
- DAWES, R. M. (1990): "The Potential Nonfalsity of the False Consensus Bias," in *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, ed. by R. M. Hogarth. University of Chicago Press, Chicago.
- DEGROOT, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York.
- DHILLON, A., AND B. LOCKWOOD (2004): "When are Plurality Rule Voting Games Dominance-Solvable?," *Games and Economic Behavior*, 46, 55–75.
- DOWNS, A. (1957): *An Economic Theory of Voting*. Harper and Row, New York.
- DROOP, H. R. (1871): "On the Political and Social Effects of Different Method of Electing Representatives," in *Papers, 1863–70*, chap. 3, pp. 469–507. Juridical Society, London.
- 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.
- EVANS, G., AND A. HEATH (1993): "A Tactical Error in the Analysis of Tactical Voting: A Response to Niemi, Whitten and Franklin," *British Journal of Political Science*, 23, 131–137.
- FARQUHARSON, R. (1969): *Theory of Voting*. Yale University Press, New Haven.
- FEDDERSEN, T. J. (1992): "A Voting Model Implying Duverger's Law and Positive Turnout," *American Journal of Political Science*, 36(4), 938–62.
- FEDDERSEN, T. J., AND W. PESENDORFER (1996): "The Swing Voter's Curse," *American Economic Review*, 86(3), 408–24.
- (1997): "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 65(5), 1029–58.
- (1998): "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review*, 92(1), 23–35.

- FEREJOHN, J. A., AND M. P. FIORINA (1974): "The Paradox of Not Voting: A Decision Theoretic Analysis," *American Political Science Review*, 68(2), 525–36.
- 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.
- (2001a): "Definition and Measurement of Tactical Voting: The Role of Rational Choice," *CREST Working Paper* 95, available from www.crest.ac.uk. Forthcoming in the *British Journal of Political Science*.
- (2001b): "Extending the Rational Voter Theory of Tactical Voting," Paper presented at the Mid-West Political Science Association Meeting in Chicago.
- FRANKLIN, M., R. G. NIEMI, AND G. WHITTEN (1994): "The Two Faces of Tactical Voting," *British Journal of Political Science*, 24, 549–557.
- GELMAN, A., J. N. KATZ, AND F. TUERLINCKX (2002): "The Mathematics and Statistics of Voting Power," *Statistical Science*, 17(4), 420–35.
- GELMAN, A., AND G. KING (1994): "A Unified Method of Evaluating Electoral Systems and Redistricting Plans," *American Journal of Political Science*, 38, 514–54.
- GOEREE, J. K., AND J. GROSSER (2003): "Costly Voting with Correlated Preferences," University of Amsterdam.
- GOLDSTEIN, I., AND A. PAUZNER (2003): "Demand Deposit Contracts and the Probability of Bank Runs," Duke University and Tel Aviv University.
- GOOD, I. J., AND L. S. MAYER (1975): "Estimating the Efficacy of a Vote," *Behavioral Science*, 20, 25–33.
- GREEN, D. P., AND I. SHAPIRO (1994): *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. Yale University Press, New Haven, CT.
- GREENBERG, J. (1979): "Consistent Majority Rules Over Compact Sets of Alternatives," *Econometrica*, 47(3), 627–36.
- 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.
- HOTELLING, H. (1929): "Stability in Competition," *Economic Journal*, 39, 41–57.
- 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.
- KRAMER, G. H. (1973): "On a Class of Equilibrium Conditions for Majority Rule," *Econometrica*, 41(2), 285–97.
- (1977): "A Dynamic Model of Political Equilibrium," *Journal of Economic Theory*, 16, 310–34.
- LANOUE, D. J., AND S. BOWLER (1992): "The Sources of Tactical Voting in British Parliamentary Elections, 1983–1987," *Political Behaviour*, 14, 141–157.
- LEDYARD, J. O. (1984): "The Pure Theory of Large Two-Candidate Elections," *Public Choice*, 44(1), 7–41.

- MARGOLIS, H. (1977): "Probability of a Tie Election," *Public Choice*, 31, 134–7.
- MARGOLIS, H. (1983): "The Banzhaf Fallacy," *American Journal of Political Science*, 27(2), 321–26.
- MCKELVEY, R., AND P. C. 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.
- (1985): "Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources," *Journal of Economic Theory*, 36, 55–85.
- MEEHL, P. E. (1977): "The Selfish Voter Paradox and the Thrown-Away Vote Argument," *American Political Science Review*, 71, 11–30.
- MORRIS, S., AND H. S. SHIN (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88(3), 587–97.
- (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics, the Eighth World Congress*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge University Press, Cambridge, UK.
- (2004): "Coordination Risk and the Price of Debt," *European Economic Review*, 48(1), 133–53.
- MORTON, R. B. (1991): "Groups in Rational Turnout Models," *American Political Science Review*, 35(3), 758–76.
- (2001): *How Elections Rule American Politics*. Norton, forthcoming.
- MULLIGAN, C. B., AND C. G. HUNTER (2003): "The Empirical Frequency of a Pivotal Vote," *Public Choice*, 116(1–2), 31–54.
- MYATT, D. P., AND S. D. FISHER (2002): "Tactical Coordination in Plurality Electoral Systems," *Oxford Review of Economic Policy*, 18(4), 504–22.
- (2003): "Everything is Uncertain and Uncertainty is Everything," University of Oxford, Department of Economics Discussion Paper.
- MYATT, D. P., H. S. SHIN, AND C. WALLACE (2002): "The Assessment: Games and Coordination," *Oxford Review of Economic Policy*, 18(4), 397–417.
- MYERSON, R. B. (1999): "Theoretical Comparisons of Electoral Systems," *European Economic Review*, 43, 671–97.
- (2000): "Large Poisson Games," *Journal of Economic Theory*, 94(1), 7–45.
- (2002): "Comparison of Scoring Rules in Poisson Voting Games," *Journal of Economic Theory*, 103, 219–51.
- MYERSON, R. B., AND R. J. WEBER (1993): "A Theory of Voting Equilibria," *American Political Science Review*, 87(1), 102–114.
- NIEMI, R. (1984): "The Problem of Strategic Voting under Approval Voting," *American Political Science Review*, 78, 952–8.
- 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.

- NIEMI, R. G., G. WHITTEN, AND M. FRANKLIN (1993): "People Who Live in Glass Houses: A Response to Evans and Heath's Critique of our Note on Tactical Voting," *British Journal of Political Science*, 23, 549–563.
- OSBORNE, M. J. (1995): "Spatial Models of Political Competition under Plurality Rule: A Survey of Some Explanations of the Number of Candidates and the Positions They Take," *Canadian Journal of Economics*, 28(2), 261–301.
- OSBORNE, M. J., AND A. RUBINSTEIN (2003): "Sample Equilibrium, with an Application to Strategic Voting," *Games and Economic Behavior*, 45, 434–41.
- 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.
- PALFREY, T. R., AND H. ROSENTHAL (1983): "A Strategic Calculus of Voting," *Public Choice*, 41, 7–53.
- (1985): "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 79, 62–78.
- 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.
- PIKETTY, T. (2000): "Voting as Communicating," *Review of Economic Studies*, 67(1), 169–91.
- RABIN, M., AND J. L. SCHRAG (1999): "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, 114(37–82).
- RIKER, W. H. (1976): "The Number of Political Parties: A Reexamination of Duverger's Law," *Comparative Politics*, 9, 753–766.
- (1982a): *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. Waveland Press, Prospect Heights, IL.
- (1982b): "The Two-Party System and Duverger's Law: An Essay on the History of Political Science," *American Political Science Review*, 76(4), 753–766.
- RIKER, W. H., AND P. C. ORDESHOOK (1968): "A Theory of the Calculus of Voting," *American Political Science Review*.
- ROSS, L., D. GREENE, AND P. HOUSE (1977): "The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes," *Journal of Personality and Social Psychology*, 13, 279–301.
- SIMON, H. A. (1954): "Bandwagon and Underdog Effects and the Possibility of Equilibrium Predictions," *Public Opinion Quarterly*, 18, 245–253.
- SIMPSON, P. B. (1969): "On Defining Areas of Voter Choice: Professor Tullock on Stable Voting," *Quarterly Journal of Economics*, 83(3), 478–90.