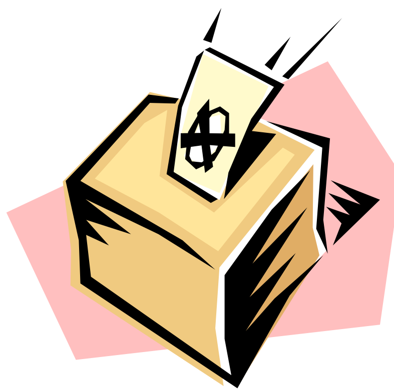
Nuffield College, University of Oxford



A New Theory of Strategic Voting

David Peter Myatt

Nuffield College and St. Catherine's College, University of Oxford

June 1999

Abstract

This is an analysis of strategic voting under the plurality rule. Existing theories predict strict bipartism, where rational voters support only two candidates: a strict interpretation of Duverger's Law. This conclusion is rejected. The new theory employs a simple model of a three-candidate election. Unlike previous models, the level of popular support for each candidate is not commonly certain. Voters form opinions of candidate support from public and private signals. Unless knowledge of constituency-wide factors is both common and precise, there is a uniquely stable non-Duvergian equilibrium, with only partial strategic voting. Strategic voting is increasing in the precision of voter beliefs, the expected strength of a leading candidate, and the expected gap between challenging candidates. The effect of information depends critically on its source. Public signals of candidate support have a far stronger effect than privately observed signals. Surprisingly, when voting decisions are based largely on private rather than public information, strategic voting is self-attenuating rather than self-reinforcing.

Acknowledgements: Stephen D. Fisher inspired this work with his extensive empirical research on tactical voting, and with many hours of conversation on the topic. Thanks are also due to David Firth, Geoff Evans, Iain McLean, Chris Wallace and seminar participants at Oxford, UCL and LSE for helpful comments.

Extended Abstract

A New Theory of Strategic Voting describes a new theory of strategic voting in plurality electoral systems. Existing theories are shown to rest on the assumed independence of voter preferences. Replacing this assumption with one of uncertain candidate support, the new theory models both the public and private information sources that enter into a voter's decision. Suprisingly, with pure private information sources, strategic voting is self-attenuating and not self-reinforcing. This leads to a rejection of strict bipartism, and a new range of comparative static results. Crucially, and contrary to existing theories, the model shows that multi-candidate support in plurality elections is perfectly consistent with rational voting. The theory is presented in a sequence of three parts.

Part 1 provides a critique of existing models. Following earlier work, it observes that the key criterion for a strategic voter is the relative likelihood of pivotal ties between the competing candidates. Existing models impose independence on the preferences of voters, yielding independence of voting decisions. Under this scenario, the relative likelihood of a tie between the two leading candidates in an election becomes infinitely greater than a tie between any other pair for large electorates, yielding a strict bipartite outcome with rational voters. Following the introduction of uncertainty over constituency-wide support, this result fails. It follows that in a decision-theoretic framework, a rational electorate gives support to all candidates. Crucially, it *only* the constituency uncertaint that matters in determining voting behaviour. Intuitively, any idiosyncratic uncertainty on the part of individual voters is averaged out in a large electorate.

Part 2 moves to a game theoretic framework. The model builds a microfoundation for constituency uncertainty in an election, by allowing for unknown constituency-wide common effects. Beliefs over such effects (and hence candidate support) are elicited from observed signals, corresponding to samples of preferences from the electorate. To separate the effects of certain support and common knowledge of expected candidate standings, the analysis restricts to public signals. Strategic voting is self-reinforcing, and with sufficient information precision, a bipartite outcome emerges. Importantly, the analysis shows that it is information precision, rather than large electorates, that can lead to bipartism. This precision can be indexed by the size of an opinion poll: Allowing the opinion poll sample size to expand leads to bipartism.

Part 3 overturns this result. It argues that, at a constituency level, information sources are likely to be privately observed. Specifying a model in which each voter observes a private signal of candidate support, it arrives at the surprising conclusion that strategic voting is self-attenuating rather than self-reinforcing. Anticipation of strategic voting by others dampens strategic incentives, hence reducing the equilibrium level of strategic switching. It follows that strategic voting is once again incomplete, and multi-candidate support emerges. With a combination of public and private information sources, a bipartite outcome requires public signals to be sufficiently precise relative to private signals.

Parts 1–3 display common comparative statics. Strategic voting is increasing in the precision of information, the homogeneity of the electorate, and the perceived gap between challenging candidates. Contrary to earlier informal analyses, there is little strategic voting in a close election. The theory shows that the lack of common knowledge of candidate standing can prevent the coordination of challenging candidates in the defeat of a disliked third party.

Contents

Part 1

Deconstructing Duverger's Law

"The whole problem with the world is that fools and fanatics are always so certain of themselves, but wiser people so full of doubts."

Bertrand Russell

1.1 Introduction

1.1.1 Duverger's Law

The tendency of plurality voting to yield a two-party system is a long-held tenet. In 1951, Maurice Duverger elevated this hypothesis to legal status. In an oft-quoted passage, he stated that the hypothesis "approaches most nearly perhaps to a true sociological law" (Duverger 1954, p. 217). This daring statement is, however, accompanied by qualifying language. As Riker (1982, p. 754) notes, Duverger's hypotheses display some ambiguity, perhaps best exhibited by the statement of Duverger's Law itself as "the simple-majority single-ballot system *favors* the two-party system" (Duverger 1954, p. 217, emphasis added). This stops short of claiming that a plurality system is sufficient for strict bipartism.

The formulation of Duverger's Law encompasses strategic interaction between both political parties and voters: The electorate responds to the stance of political parties via vote choice, and candidates act with reference to both past and future voting decisions. It is clear, however, that the influence of the voting mechanism must stem from its effect on voting behaviour.[1] Analysis, therefore, must include investigation of voting under the plurality rule, and the associated district level bipartism. Consideration of strategic voting and the wasted vote phenomenon

[1] Note that this may be *anticipated* rather than *actual* voter behaviour. There may only be two political parties, since an entrant *anticipates* performing poorly under the electoral system. It is still the effect of the voting mechanism on voter behaviour that drives this — even though it is a hypothesised election rather than an actual one.

follows — instrumentally rational voters may choose to vote for someone other than their preferred candidate. As Droop (1871) and many followers have observed, a voter may notice that their preferred candidate has little chance of winning the election. Mindful of "wasting their vote", an individual may switch their vote to one of the leading contenders. Such strategic behaviour leads to erosion of support for lower placed candidates, favouring the leading pair.

Formal analyses of strategic voting have led to a rather draconian amendment to Duverger's Law. These models are surveyed by Gary Cox (1997), and claim that, under the plurality rule, only two candidates receive votes from an instrumentally rational electorate. Palfrey (1989) is a case in point — his paper presents a "mathematical proof" to enforce this strict version of Duverger's Law:

> " ... with instrumentally rational voters and fulfilled expectations, multicandidate contests under plurality rule should result in only two candidates getting any votes."

Palfrey examines the stable Bayesian Nash equilibria of his model, and finds that the support for any third candidate vanishes completely as the electorate grows large. In contrast to this prediction many plurality voting systems, including those of the UK and India, exhibit multi-candidate support.[2] Cox (1997) and others rightly recognise that their voting models are stylised abstractions, and hence one should not expect to observe their exact conclusions. In the light of Palfrey's claim, it is tempting to attribute the absence of strict bipartite election outcomes to the lack of rationality of voters, or to miscoordination on their part.

1.1.2 A New Theory of Strategic Voting

The new theory rejects this assertion. It observes that the assumed independence of voter preferences drives existing models. In response, it develops a new model in which voters are uncertain of constituency wide-preferences, and carefully models both public and private information sources upon which individuals condition their voting decisions. Arguing that privately-observed information sources are of key importance at a district level, the analysis shows that in the private information case strategic voting is self-attenuating rather than self-reinforcing. This leads to non-Duvergerian conclusions, with multi-candidate support, and new comparative statics. Summarising, the observation of support for multiple candidates in plurality elections is perfectly consistent with rational behaviour on the part of individual voters.

The formulation of the theory begins here. This part analyses the decision-theoretic case, and explores the implications of weakening the independence of voter preferences. Part 2 presents

[2] Analysis of strategic voting in Britain is provided by Johnston and Pattie (1991), Lanoue and Bowler (1992) and Niemi, Whitten and Franklin (1992) *inter alia*. For an analysis of the Indian case, see Riker (1976). He hypothesises that reduced strategic voting is due to the presence of a clear Condorcet winner, against which strategic voting is futile. This is not supported by the analysis of the new theory. A full application to this case is left to future work.

a fully game-theoretic model, in which information pertaining to constituency-wide preferences is obtained from a public source. It demonstrates that sufficient precision in public signals is required in order to obtain a stable Duvergerian outcome. Finally, Part 3 extends the model to include private signals. The analysis shows that strategic voting may exhibit negative feedback, and reaches a non-Duvergerian conclusion.

1.1.3 Voting and Preference Independence

The argument begins here in Part 1 by noting that a key assumption drives Duvergerian results — the authors cited here all specify that the preferences of voters are drawn *independently* from a *commonly known* distribution. Given a symmetric Bayesian strategy profile, independence of preferences leads to the independence of voting decisions. In particular, the leading pair of candidates in an plurality election is then commonly known. Cox (1997, p. 78) himself recognises this fact, noting that:

> "[A condition] necessary to generate pure local bipartism is that the identity of trailing and front-running candidates is common knowledge."

He hypothesises that failure of this condition is akin to miscoordination in the Battle of the Sexes. A rather more subtle mechanism is at work, however. To see this, and following Hoffman (1982), Palfrey (1989), Myerson and Weber (1993) and others, note that in a plurality mechanism it is not the voting probabilities *per se* that are of interest to a voter. Indeed, a single individual can only influence the outcome of an election in the event of a tie for the lead. A casting vote may then determine the winning candidate. Of interest, therefore, is the relative likelihood of ties between different pairs of candidates. With independent voting probabilities, the "pivotal" probability of a tie between the two leading candidates becomes infinitely greater than a tie between any other pair, as the electorate grows large. It follows that a rational voter chooses one of the two leading candidates, and a Duvergerian outcome obtains.

If preference realisations are not independent and voting strategies are non-degenerate, then this result fails. The likelihood ratio of a tie between the two leading candidates versus a tie between any other pair converges to a finite value as the electorate grows large. Any voter with a sufficiently strong preference for a win by a third candidate will not vote strategically — a non-Duvergerian outcome.

The analysis explores this observation in more detail. A simple model of a three candidate election is constructed. The setting will be familiar to readers of Fey (1997). There are two types of voter. The first prefer the first candidate. The second type prefer both of the remaining candidates to the first. The model focuses on strategic switching between the latter candidates. Importantly, and in contrast to earlier analyses, the constituency-wide support for these latter candidates is uncertain. Constructing the optimal voting rule, the key statistic is the pivotal log-likelihood ratio of a tie between each of the two challenging candidates and the first candidate.

As the constituency size grows large, this converges to a finite value, and hence there is a finite incentive for a decision-theoretic strategic voter to switch candidates. It follows that strategic voting is incomplete.

The critical observation is that it is *only* the uncertainty over constituency-wide support that drives strategic voting. To see this, notice that uncertainty over individual preferences and hence decisions is a necessary component of strategic voting. If the outcome of the election is certain, then (except in the case of a tie) a single vote cannot influence the outcome. Uncertainty over preferences (and hence the election outcome) results in the possibility (however remote) of a tied outcome, and hence a rôle for an individual voter. Consider the possible sources of such uncertainty. First, fixing the distribution of voter preferences, the decision of each individual is uncertain, since each voter is a single draw from such a distribution. This is *idiosyncratic* uncertainty. Secondly, a typical voter may be uncertain of the constituency-wide distribution from which preferences are drawn. This is *constituency* uncertainty. The Cox-Palfrey model assumes that the preference distribution is known. It follows that their results are entirely driven by idiosyncratic uncertainty. By contrast, when both sources of uncertainty are present, it is *only* constituency uncertainty that matters. The intuition is clear. In a large constituency, the idiosyncrasies of individuals are averaged out. Constituency uncertainty, however, is common to all and hence not averaged out by the Law of Large Numbers. It follows that the Cox-Palfrey framework is driven by the wrong factors.

Many of the ideas offered both here and in the subsequent parts have been explored by previous authors. As Riker notes, analysis of the "wasted vote" phenomenon dates back to Droop (1871), with more recent interest stemming from Duverger (1954) and Farquharson (1969). The main antecedents to this work, however, are the McKelvey-Ordeshook (1972) decision-theoretic and Cox (1994), Palfrey (1989) and Myerson-Weber (1993) game-theoretic models of strategic voting. Cox anticipates that the strict bipartite outcomes are likely to fail with a "lack of public information about voter preferences and vote intentions", and the new theory provides a formalisation of this prediction.

1.1.4 Guide

The argument is formalised in the following sections. A simple model of a three-candidate election is constructed in Section 1.2. All individuals save one are assumed to vote straightforwardly. Section 1.3 considers the optimal voting decision of the lone strategic voter, and elicits the key rôle of the pivotal log-likelihood ratio. The behaviour of this statistic is illustrated in Section 1.4 with a motivating example. The main contribution of this part is in Section 1.5, with an analysis of the asymptotic behaviour of this statistic. The limiting behaviour is illustrated in Section 1.6. Section 1.7 concludes the preliminary argument, as a precursor to the game-theoretic extensions of Parts 2 and 3.

1.2 Model

1.2.1 The Election

There are three candidates in a single-seat district election, indexed by $j \in \{1, 2, 3\}$. Policy platforms have been set, and hence voting will take place without any candidate action. There are $N+1$ members of the electorate, indexed by $i \in \{0, 1, \dots, N\}$, where $N = (1+\gamma_1)n$. Voter 0 is a focal voter, and will consider the behaviour of the remaining N voters to assess the likelihood of the various electoral outcomes. Realised vote totals for the N voters $i > 0$, for each of the three parties will be denoted $\{x_j\}_{j=1}^3$, so that $\sum_{j=1}^3 x_j = N = (1+\gamma_1)n$ — hence each individual casts a single vote. The candidate receiving the largest number of votes wins the election — a plurality mechanism.

1.2.2 Voter Behaviour

The argument will focus on tactical vote switches between two candidates. Envision a situation in which voters prefer either the first candidate, or alternatively both of the remaining candidates over the first candidate. With this in mind, it is assumed that a fixed fraction $\gamma_1/(1+\gamma_1)$ of the N electors $i > 0$ will vote for candidate 1, so that $x_1 \equiv \gamma_1 n$. To obtain a model of interest, the following assumption is upheld.

Assumption 1. *The fraction γ_1 satisfies $1/2 < \gamma_1 \leq 1$.*

Notice that $\gamma_1 > 1$ yields a certain win for candidate 1. $\gamma_1 > 1/2$ ensures that any close result must be between candidate 1 and candidate $j \in \{2, 3\}$. Throughout the analysis, it is implicitly assumed that $\gamma_1 n$ is integer valued. This is without loss of generality, since all results continue to hold when $x_1 = \lceil \gamma_1 n \rceil$ or $x_1 = \lfloor \gamma_1 n \rfloor$.

The model is decision-theoretic, and hence the N voters $i > 0$ act straightforwardly, and vote for their preferred candidate. An individual i receives a payoff u_{ij} if candidate j wins the election. It follows that $x_1 = \gamma_1 n$ individuals satisfy $u_1 > \max\{u_{i2}, u_{i3}\}$. The remaining n individuals (and also the focal individual $i = 0$) satisfy $\min\{u_{i2}, u_{i3}\} > u_{i1}$ — candidate 1 is their least preferred option. It follows that $x_2 + x_3 = n$. The focal voter $i = 0$ is short-term instrumentally rational: This voter's decision is determined solely by its anticipated influence on the outcome of the election. It follows that such a voter must support either candidate 2 or 3. The model is designed to focus on strategic vote switches between these two options. With this in mind, the following notation is employed.

Definition 1. *Henceforth, all references to "a voter", "voters" or the "electorate" will refer to the $n+1$ supporting candidates 2 and 3.*

Since $x_1 = \gamma_1 n$, it follows that a fraction γ_1 of these n voters must successfully coordinate on either candidate 2 or 3 in order to defeat candidate 1. Voting is thus a coordination game between the $n+1$ individuals. Whereas this chapter will consider only the optimal response of the focal voter $i = 0$, Chapters 2 and 3 take a fully game-theoretic stance and allow strategic behaviour throughout the electorate.

It remains to specify the distribution of preferences within the $n + 1$ individuals satisfying $\min\{u_{i2}, u_{i3}\} > u_{i1}$. Denote by p the probability that a voter prefers candidate 2, so that:

$$p = \Pr[u_{i2} \geq u_{i3} \mid \min\{u_{i2}, u_{i3}\} > u_{i1}]$$

Fixing p, the model thus far is equivalent to those of previous authors, including Cox (1994) and Fey (1997). In particular, voter preferences (and hence truthful voting decisions) are independent. It follows that the expected vote share of a candidate is known. Applying the Law of Large Numbers, the realised vote shares converge in probability to the expected vote shares for large constituencies. The independence assumption is removed by allowing p to be uncertain.

Assumption 2. *The focal voter $i = 0$ is uncertain of p. Conditional on any information available, p takes values in $(0, 1)$ with continuous density $f(p) > 0$, where $f_H = \sup_p f(p) < \infty$.*

Uncertainty over p leads to interdependence of voting decisions. The decision of one individual yields an updated posterior over p, and hence changes the probability of subsequent decisions. This is made clear in the example below.

1.2.3 Parametric Example — The Beta-Binomial

The Beta distribution yields a convenient conjugate prior for the binomial distribution — see, for instance, Casella and Berger (1990). The Beta has density:

$$f(p) = \frac{\Gamma(\beta_2 + \beta_3)}{\Gamma(\beta_2)\Gamma(\beta_3)} p^{\beta_2 - 1}(1 - p)^{\beta_3 - 1}$$

This yields $\gamma_2 = \mathrm{E}[p] = \beta_2/(\beta_2 + \beta_3)$. For convenience, fix $\beta_2 = m\gamma_2$ and $\beta_3 = m(1 - \gamma_2)$. This fixes the mean at γ_2, and the parameter m may be used to vary the precision of beliefs. The associated density becomes:

$$f(p) = \frac{\Gamma(m)}{\Gamma(m\gamma_2)\Gamma(m(1 - \gamma_2))} p^{m\gamma_2 - 1}(1 - p)^{m(1-\gamma_2)-1} \tag{1.1}$$

To check the precision interpretation of m, calculate:

$$\frac{1}{\mathrm{var}[p]} = \frac{(\beta_2 + \beta_3)^2(\beta_2 + \beta_3 + 1)}{\beta_2\beta_3} = \frac{m + 1}{\gamma_2(1 - \gamma_2)}$$

The distribution is illustrated for $\gamma_2 = 32/62$ and a variety of m in Figure 1.1. As the precision m increases, the distribution concentrates all mass around $p = \gamma_2$, and hence $m \to \infty$ corresponds

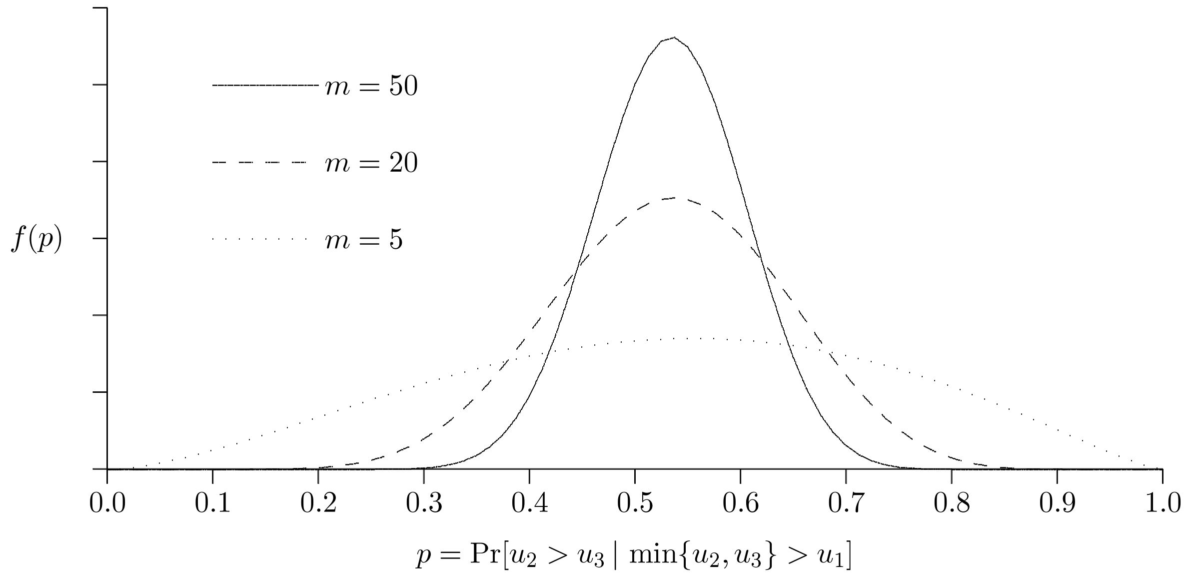


Figure 1.1: Beta Prior for Various m

to the case of certain constituency-wide support. The particular parametric choice of $\gamma_2 = 32/62$ corresponds to the motivating example of Section 1.4.

1.3 Voting Behaviour

A vote can only influence the outcome of an election if it is *pivotal*. A pivotal situation arises if there is a tie for the lead. If this occurs, a single casting vote may swing the outcome in favour of one of the candidates. More generally, a voter may also influence the outcome if the votes cast for the leading candidates differ by a single vote. The voter may then ensure the win of one candidate, or alternatively force a draw. For this latter case, the tie-break rule adopted in the election comes into play.

The characteristics of the optimal voting rule are familiar. Recall that Assumption 1 stipulated $\gamma_1 > 1/2$, so that $x_1 = \gamma_1 n > N/3$. It is clear that any tie for the lead must involve candidate 1. The possibilities are thus $x_1 = x_2 > x_3$ and $x_1 = x_3 > x_3$. Near ties, for instance $x_1 = x_2 + 1$, are also relevant. Denote the "pivotal" probabilities, conditional on the information of the focal voter $i = 0$ by q_2 and q_3 respectively:

$$q_j = \Pr[\gamma_1 n - 1 \leq x_j \leq \gamma_1 n] \text{ for } j \in \{2, 3\}$$

These are the *pivotal probabilities* of candidates 2 and 3 lying in tied (or near-tied) contention for the lead with candidate 1.

Lemma 1. *A typical individual votes for party 2 if and only if* $q_2(u_2 - u_1) \geq q_3(u_3 - u_1)$.

If there is a tie between candidates 1 and 2, the voter can shift the result from candidate 2, so gaining $u_2 - u_1$. The voter loses the opportunity to make the casting vote in a tie between candidates 1 and 3, sacrificing $u_3 - u_1$. The voting rule trivially follows.[3] This sketch neglects the outcomes from a dead tie. A fuller proof follows.

Proof of Lemma 1: Recall that the restriction is to voters i satisfying $\min\{u_{i2}, u_{i3}\} > u_{i1}$. A voter will never find it optimal to vote for candidate 1, and hence will choose to vote for one of candidates $j \in \{2, 3\}$. If $x_j > \gamma_1 n$, then candidate j will win without the support or otherwise of voter i, and hence the vote has no influence. Similarly, if $\max\{x_j\} < \gamma_1 n - 1$, the additional vote does not influence the election. If, however, for some $j \in \{2, 3\}$ there is a tie $x_j = \gamma_1 n$, then an additional vote for candidate j will yield a win. Similarly, if $x_j = \gamma_1 n - 1$, then an additional vote for candidate j will force a tie. Impose the assumption that that ties are broken at random, corresponding to British electoral convention. Then the potential outcomes and payoffs from voting for candidates $j \in \{2, 3\}$ are as follows:

Event	$u(\text{Vote } 2)$	$u(\text{Vote } 3)$	$u(\text{Vote } 2) - u(\text{Vote } 3)$
$x_2 = \gamma_1 n - 1$	$\frac{u_{i2}+u_{i1}}{2}$	u_{i1}	$\frac{u_{i2}-u_{i1}}{2}$
$x_2 = \gamma_1 n$	u_{i2}	$\frac{u_{i2}+u_{i1}}{2}$	$\frac{u_{i2}-u_{i1}}{2}$
$x_3 = \gamma_1 n - 1$	u_{i1}	$\frac{u_{i3}+u_{i1}}{2}$	$-\frac{u_{i3}-u_{i1}}{2}$
$x_3 = \gamma_1 n$	$\frac{u_{i3}+u_{i1}}{2}$	u_{i3}	$-\frac{u_{i3}-u_{i1}}{2}$

Assembling these payoffs, the optimal rule is clearly:

$$\text{Vote } 2 \quad \Leftrightarrow \quad \Pr[\gamma_1 n \geq x_2 \geq \gamma_1 n - 1]\frac{u_{i2} - u_{i1}}{2} \geq \Pr[\gamma_1 n \geq x_3 \geq \gamma_1 n - 1]\frac{u_{i3} - u_{i1}}{2}$$

Recalling the definition of q_j, this is the desired result. □

A further implicit assumption of Lemma 1 is that $\min\{q_j\} > 0$. Such a condition rules out Duvergerian outcomes, where all n voters support only one of the candidates $j \in \{2, 3\}$. In this case, unless $\gamma_1 = 1$, the tie probabilities are exactly zero. Within such a scenario, a voter is exactly indifferent between the two options — for positive q_j this is (generically) never the case. In the decision-theoretic analysis of this part, the focal voter acts on the assumption of straightforward voting by the remainder of the electorate, and this case will never arise. The game-theoretic analysis of Parts 2 and 3 must address this issue, however, when considering the possibility of Duvergerian equilibria. For positive q_j, the following notation is introduced.

Definition 2. *The (perceived) log-likelihood ratio of pivotal probabilities is* $\lambda = \log(q_2/q_3)$.

Using this notation and Lemma 1, the optimal voting rule becomes:

$$\text{Vote } 2 \quad \Leftrightarrow \quad \log \frac{u_{i2} - u_{i1}}{u_{i3} - u_{i1}} + \lambda \geq 0$$

[3] An assumption implicit to Lemma 1 is that an indifferent voter casts their vote for candidate 2. This is made without loss of generality.

Hence the key statistic of interest to a voter is the log-likelihood ratio λ. The tradeoff is clear. A strategic voter balances the relative likelihood of each candidate j contending for the lead against the relative preference for a win by each of the two candidates. Notice also that λ provides an exact measure of the incentive for a voter to act strategically. In particular, if an individual votes straightforwardly, then they vote for candidate 2 whenever $\log(u_2 - u_1) \geq \log(u_3 - u_1)$. This is equivalent to using the optimal voting rule with $\lambda = 0$.

1.4 Illustration

1.4.1 Strategic Voting with Certain Candidate Shares

The key assumptions behind the Duvergerian conclusions are best understood with reference to the following example. To avoid the numeric reference of candidates, the names of the three main British political parties have been employed.[4] Needless to say, the example is pedagogical and hence not designed to represent any particular constituency.

Example 1. *In a three candidate plurality election, the electors cast their votes truthfully. Voter preferences are independent. The expected vote shares are:*

	Candidate	***Share***
1	*Conservative*	*0.38*
2	*Labour*	*0.32*
3	*Liberal Democrat*	*0.30*

In Example 1, the Conservative candidate is identified as the leader. The trailing Labour and Liberal Democrat candidates are close, but with Labour slightly ahead. Notice that the assumption of independent voting probabilities is retained. Equivalently, the voting probabilities are known with certainty. Employing the plurality rule, the winner is the candidate receiving the most votes. Straightforward voting leads to a Conservative win.

Consider the introduction of an instrumentally rational voter. As is standard, such an individual cares only about their vote insofar as it effects the outcome of the election. Their optimal choice may involve switching their vote away from their preferred candidate. Suppose that the electorate is of size N. Cox (1997, p. 75) indexes the candidates by j, and denotes the

[4] In the British General Election of 1997, the Labour and Liberal Democrat opposition parties challenged the incumbent Conservative Party. Curtice and Steed (1997) note that in seats held by the Conservatives, voters switched between each of the opposition parties in an attempt to dislodge the unpopular incumbent. Fixing the Conservative vote share and examining switching between Labour and the Liberal Democrats provides a reasonable specification. See Butler and Kavanagh (1997) for more details of the British general election. It is worthwhile noting that in British general elections strategic switching has occurred between *all* political parties. It follows that the model considered here is not exact — rather, it illustrates the key factors influencing the decision of a strategic voter.

associated voting probabilities by π_j, where the leading candidates are $j \in \{1, 2\}$. He explains the standard strategic voting logic as follows:

> "If $0 < \pi_j < \pi_2$ then candidate j is virtually sure to lose for sufficiently large N, and voting for the most palatable of the candidates most likely to be tied for first yields a higher expected utility than voting for j."

The first part of this statement is empty — for sufficiently large N, candidate 2 is also "virtually sure to lose". Indeed, in Example 1, with a large electorate, the Conservative candidate will win with near certainty — this is a simple consequence of the Law of Large Numbers. The key criterion (as Cox correctly recognises) is the probability of a candidate being in tied contention for the lead — a *pivotal* outcome. Although this is a remote possibility, it is the only instance in which an extra vote can swing the outcome of the election. This *pivotal probability* becomes vanishingly small as N grows large. An instrumentally rational voter, however, is only interested in the *relative* likelihood of the various possible pivotal possibilities *given* that a pivotal outcome actually occurs — if a pivotal outcome does not occur, a single vote carries no weight and hence such situations are irrelevant. These are *conditional pivotal probabilities* — for instance, the probability of a Conservative-Labour tie[5], given that a tie between any two candidates occurs.

To sharpen the analysis, place Example 1 into the model of Section 1.2. Fix the Conservative vote at $x_1 = \gamma_1 n$, and the total electorate at $N = (1 + \gamma_1)n$, where $\gamma_1 = 0.38/0.62$. Inspecting the specification of Example 1, it is clear that any pivotal outcome must involve the Conservative candidate. Moreover, one might intuitively expect that a Con-Lab tie for the lead is more likely than a Con-LibDem tie. With this in mind, a Liberal Democrat supporter might switch to Labour in an attempt to prevent a Conservative win — or indeed vice versa. In a decision theoretic setting, one can thus expect at least *some* strategic voting. This leaves open the question of whether this strategic voting will be complete. Clearly, a Liberal Democrat supporter will need to balance a preference for their first choice against the relative likelihood of the different tie probabilities. Since the Labour and Liberal Democrat candidates are "close" in Example 1, one might expect the conditional pivotal probabilities to be close, and thus a sufficiently committed Liberal Democrat may find it optimal to vote truthfully. Intuition, however, is insufficient to reach this conclusion — calculation of the relevant probabilities is necessary.

Figure 1.2 presents the conditional pivotal probabilities for Example 1. They are calculated for a range of electorate sizes. Even for small electorates it is far more likely that a Con-Lab tie is observed than any other tie for the lead. In fact, for $n = 500$, a Con-Lab tie is approximately 1500 times more likely than a Con-LibDem tie. For larger n, any tie is virtually certain to be between the Conservative and Labour candidates.

---

[5] Henceforth tying pairs will be referred to by the abbreviations Con-Lab, Lab-LibDem and Con-LibDem.

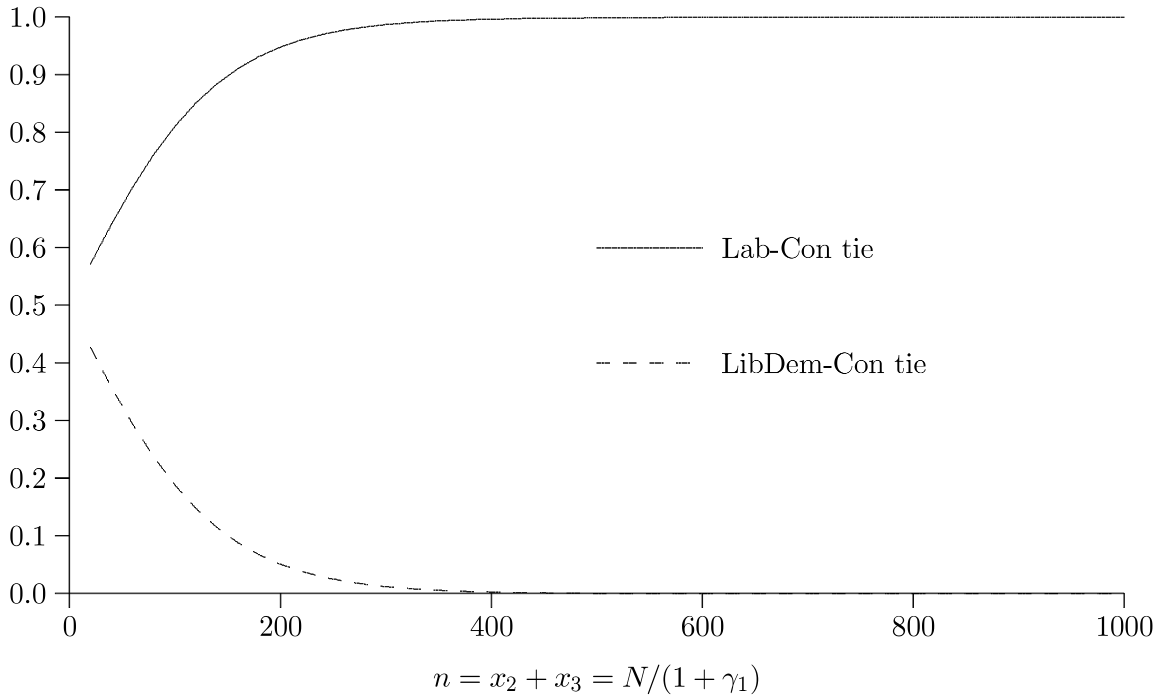


Figure 1.2: Pivotal Probabilities for Certain $p = \gamma_2 = 0.32/0.62$

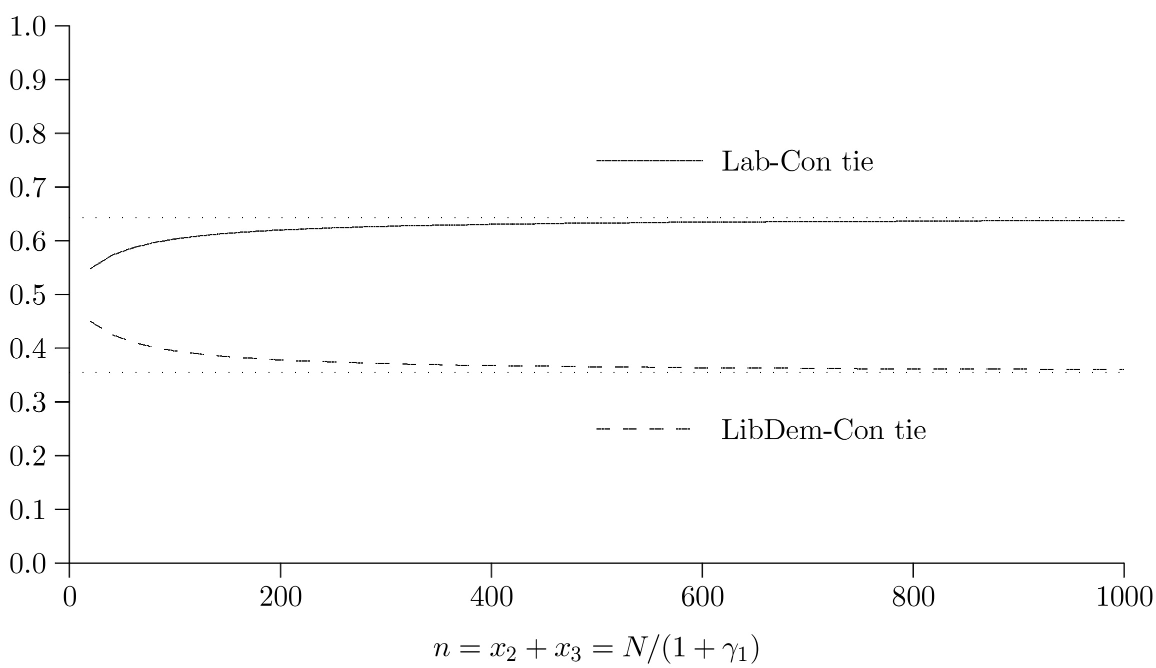


Figure 1.3: Pivotal Probabilities for Uncertain p

This example clearly illustrates the driving force behind the strict version of Duverger's Law. However close the trailing parties are, as the electorate size grows large the conditional pivotal probability of a tie between the leading two parties approaches 1. Almost all instrumentally rational voters will thus switch their vote to their preferred candidate from the two leaders. Notice that game theoretic reasoning is unnecessary — a Duvergerian outcome obtains without the need for a voter to consider the strategic actions of others. Of course, if they do consider this, it only serves to reinforce the effect. Essentially, with a large electorate there is no uncertainty remaining in the mind of a voter. All of the idiosyncrasy of individual electors is averaged out by the electorate size.

1.4.2 Strategic Voting with Uncertain Candidate Shares

The conclusions of Example 1 are both unintuitive and unattractive. For an unboundedly large electorate, the model implies that voters are certain that the Conservative and Labour candidates are the leaders, and moreover are certain that the Liberal Democrat candidate will never be in contention for the lead. Certainty over voting probabilities may be a simplifying assumption, but it is the feature that drives all the results. This observation is illuminated with reference to the following extension of Example 1.

Example 2. *Retaining the framework of Example 1, in a three candidate plurality election, the electors cast their vote truthfully. Conditional on knowledge of constituency-wide support, voters are independent. As before, the Conservative share is fixed at 0.38. There is uncertainty, however, over the probability of a voter supporting the Labour candidate, represented by a Beta prior over* p*, with* $E[p] = \gamma_2 = 0.32/0.62$*.*

In Example 2, the assumption of certain voting probabilities is relaxed. Whereas the Conservative vote share is known, there is uncertainty over the Labour (and hence Liberal Democrat) voting probabilities. The Beta distribution of p is that described in Section 1.2.3, with $m = 40$, $\beta_2 = m\gamma_2$ and $\beta_3 = m(1 - \gamma_2)$. The expected value of this distribution is $\gamma_2 = 0.32/0.62$. A voter in isolation is believed to support Labour with this probability — for a single individual, the voting probabilities of Example 1 are retained. This does *not* extend to multiple voters. If a voter is known to support Labour, this is a signal of a higher Labour support probability. An additional voter would be expected to support Labour with probability greater than γ_2. It follows that the preferences of electors are no longer independent. This invalidates the application of the Law of Large Numbers. As a consequence, the outcome of an election (with truthful behaviour) is no longer certain in large electorates. The focal voter believes that the Labour candidate is the most likely challenger, but is not certain that this is so.

Once again, it is not the voting probabilities that determine the behaviour of an instrumentally rational voter, but rather the conditional pivotal probabilities. Figure 1.3 plots these probabilities for Example 2. Notice that the conditional probability of a Con-LibDem tie does *not* fall to zero. Of course, a Con-Lab tie is more likely than a Con-LibDem tie, and hence there will be at least some incentive for strategic voting from the Liberal Democrat to Labour. Since the possibility of LibDem contention is still open, however, a sufficiently partisan Liberal Democrat supporter will find it optimal to vote truthfully. Strategic voting is not complete. It follows that, in a decision-theoretic environment, the strict version of Duverger's Law fails.

Figure 1.3 suggests that the conditional pivot probabilities tend to some finite limiting value as the constituency size grows large. Moreover, this limit is approached for relatively small constituency sizes. The analysis of Section 1.5 investigates.

1.5 Pivotal Properties

Section 1.3 highlighted the pivotal log-likelihood ratio $\lambda = \log(q_1/q_2)$ as the clear statistic of interest. Moreover, the second illustrative example of Section 1.4 suggested a finite limiting value for λ for large constituencies. This section provides a formal investigation of its limiting properties. The first step is the evaluation of the pivotal probabilities $\{q_j\}$.

Lemma 2. *The pivotal probability q_2, conditional on the information of a typical voter is:*

$$q_2 = \frac{n!}{(\gamma_1 n)!((1-\gamma_1)n)!} \int_0^1 \left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n \left\{1 + \frac{1-p}{p} \frac{\gamma_1}{1-\gamma_1+1/n}\right\} f(p)\, dp \tag{1.2}$$

A symmetric expression holds for q_3.

Proof. Simple application of the binomial distribution — see Section 1.8. □

Notice that the pivotal probabilities q_j vanish as n grows large — with a large electorate, the probability of observing a tied outcome becomes vanishingly small. Rather than these pivotal probabilities, it is the log-likelihood ratio λ that is of interest. With uncertainty over p, this converges to an attractive expression. The following proposition is the central contribution of this part, and will be key in the game-theoretic analyses of Parts 2 and 3.

Proposition 1. *Allowing the electorate size to grow large:*

$$\lim_{n\to\infty} \log \frac{q_2}{q_3} = \log \frac{f(\gamma_1)}{f(1-\gamma_1)}$$

Examining Equation (1.2), notice that the integrand vanishes to zero due to the leading term $[p^{\gamma_1}(1-p)^{1-\gamma_1}]^n$. As n grows, this expression becomes increasingly more peaked around its maximum. This maximum occurs at $p = \gamma_1$, and hence only density at this point contributes any weight. By contrast, the integrand of q_3 is maximised at $1-\gamma_1$. A formal proof draws upon from the following lemmata.

Lemma 3. *For large constituencies, there is equal probability of a tie or near-tie:*

$$\lim_{n\to\infty} \frac{\Pr[x_j = \gamma_1 n]}{\Pr[x_j = \gamma_1 n - 1]} = 1$$

Proof. The formal proof mimics that of Lemma 4 — see Section 1.8 □

Lemma 4. *For large constituencies, the likelihood ratio of exact ties tends to* $f(\gamma_1)/f(1-\gamma_1)$*:*

$$\lim_{n\to\infty}\left\{\frac{\Pr[x_2=\gamma_1 n]}{\Pr[x_3=\gamma_1 n]}\right\}=\frac{f(\gamma_1)}{f(1-\gamma_1)}$$

Proof. Begin by recalling the formulae for $\Pr[x_2=\gamma_1 n]$. It follows that:

$$\frac{\Pr[x_2=\gamma_1 n]}{\Pr[x_3=\gamma_1 n]}=\frac{\int_0^1\left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n f(p)dp}{\int_0^1\left[(1-p)^{\gamma_1}p^{1-\gamma_1}\right]^n f(p)dp}=\frac{\int_0^1\left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n f(p)dp}{\int_0^1\left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n f(1-p)dp}$$

where the second inequality follows from a simple change of variables in the denominator. Introduce the notation $G(p)$:

$$G(p)=\frac{p^{\gamma_1}(1-p)^{1-\gamma_1}}{\gamma_1^{\gamma_1}(1-\gamma_1)^{1-\gamma_1}}\quad\Rightarrow\quad\frac{\Pr[x_2=\gamma_1 n]}{\Pr[x_3=\gamma_1 n]}=\frac{\int_0^1 G(p)^n f(p)dp}{\int_0^1 G(p)^n f(1-p)dp}\tag{1.3}$$

Notice that $G(p)$ is increasing from $G(0)=0$, attaining a maximum of $G(\gamma_1)=1$ at $p=\gamma_1$, and then declining back to $G(1)=0$. Next fix $1/2>\epsilon>0$. For notational convenience, define:

$$\begin{aligned}f_{L,\epsilon}(x)&=\inf_{x-\epsilon\le p\le x+\epsilon}f(p)\\ f_{H,\epsilon}(x)&=\sup_{x-\epsilon\le p\le x+\epsilon}f(p)\\ f_H&=\sup_{0\le p\le 1}f(p)\end{aligned}$$

Here the assumption that $f(p)$ is bounded above has been used. Employing this notation, formulate an upper bound for the ratio in Equation (1.3):

$$\frac{\Pr[x_2=\gamma_1 n]}{\Pr[x_3=\gamma_1 n]}\le\frac{f_{H,\epsilon}(\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon}G(p)^n dp+f_H\left[\int_0^{\gamma_1-\epsilon}G(p)^n dp+\int_{\gamma_1+\epsilon}^1 G(p)^n dp\right]}{f_{L,\epsilon}(1-\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon}G(p)^n dp}\tag{1.4}$$

The right hand side of Equation (1.4) has three terms. These will be considered in turn. First:

$$\frac{f_{H,\epsilon}(\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon}G(p)^n dp}{f_{L,\epsilon}(1-\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon}G(p)^n dp}=\frac{f_{H,\epsilon}(\gamma_1)}{f_{L,\epsilon}(1-\gamma_1)}$$

Next consider the second term. $G(p)$ is increasing from $\gamma_1-\epsilon$ to γ_1 and hence:

$$\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon}G(p)^n dp\ge\int_{\gamma_1-\epsilon}^{\gamma_1}G(p)^n dp\ge\epsilon G(\gamma_1-\epsilon)^n$$

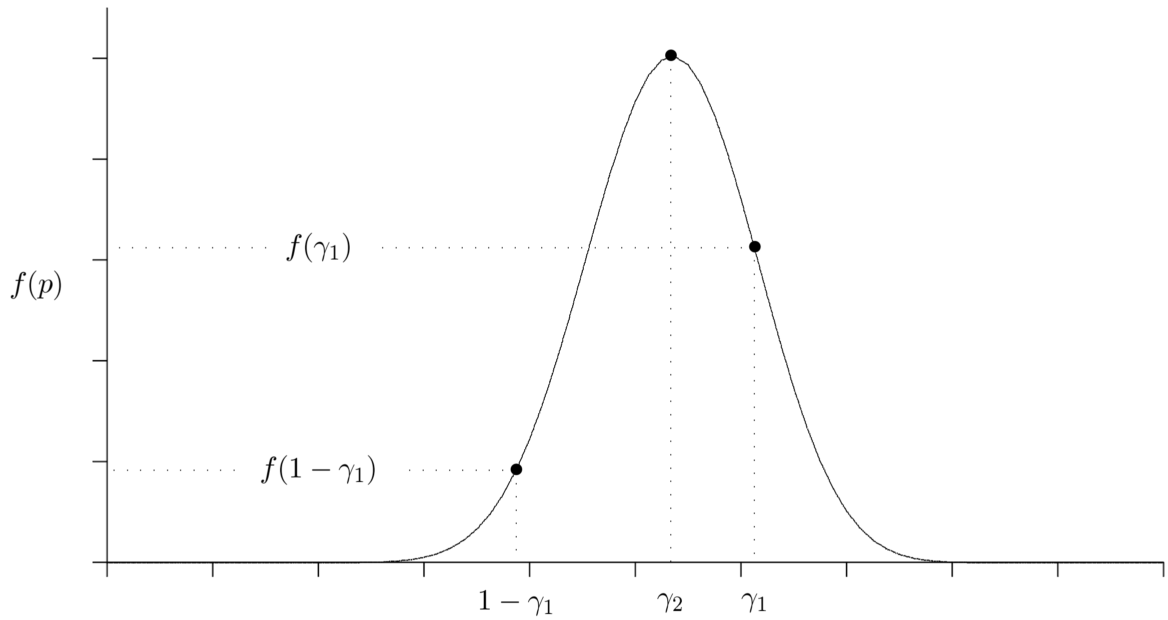


Figure 1.4: Beta Prior Over p for $m = 40$

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

$$\frac{\int_0^{\gamma_1-\epsilon} G(p)^n dp}{\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon} G(p)^n dp} \leq \frac{1}{\epsilon}\int_0^{\gamma_1-\epsilon}\left[\frac{G(p)}{G(\gamma_1-\epsilon)}\right]^n dp \longrightarrow 0$$

which holds since $G(p) < G(\gamma_1 - \epsilon)$ for all $p < \gamma_1 - \epsilon$. An identical argument ensures that the third term vanishes. Conclude from this that:

$$\lim_{n\to\infty}\frac{\Pr[x_2 = \gamma_1 n]}{\Pr[x_3 = \gamma_1 n]} \leq \frac{f_{H,\epsilon}(\gamma_1)}{f_{L,\epsilon}(1-\gamma_1)}$$

Notice now that ϵ may be chosen arbitrarily small. It follows that:

$$\lim_{n\to\infty}\frac{\Pr[x_2 = \gamma_1 n]}{\Pr[x_3 = \gamma_1 n]} \leq \lim_{\epsilon\to 0}\frac{f_{H,\epsilon}(\gamma_1)}{f_{L,\epsilon}(1-\gamma_1)} = \frac{f(\gamma_1)}{f(1-\gamma_1)}$$

A symmetric procedure bounds the limit below, and the result obtains. □

Combine Lemmata 3 and 4 to obtain the proof of Lemma 2.

Proof of Proposition 1: Employing Lemmata 4 and 3, it is clear that:

$$\frac{q_2}{q_3} = \frac{\Pr[x_2 = \gamma_1 n](1 + \Pr[x_2 = \gamma_1 n - 1]/\Pr[x_2 = \gamma_1 n])}{\Pr[x_3 = \gamma_1 n](1 + \Pr[x_3 = \gamma_1 n - 1]/\Pr[x_3 = \gamma_1 n])} \to \frac{f(\gamma_1)}{f(1-\gamma_1)}$$

This proves the desired result. □

A number of observations are pertinent. First, notice that the pivotal log-likelihood ratio λ converges to a *finite* value as the constituency grows large. Second — and perhaps more importantly — Proposition 1 implies that strategic voting is entirely driven by uncertainty over the probability p. It is the uncertainty over the constituency-wide support of the candidates, rather than uncertainty over the idiosyncratic behaviour of a particular individual that drives strategic voting. Figure 1.4 illustrates this point for Example 2. For large n, λ converges to the log-likelihood ratio of an exact tie in constituency wide support of Lab-Con ($p = \gamma_1$) versus LibDem-Con ($p = 1 - \gamma_1$).

This is a stark contrast to the Cox-Palfrey case. Suppose that constituency-wide support of all candidates is known, so that $p = \gamma_2$. Modifying Lemma 2 appropriately yields:

$$q_2 = \frac{n!}{(\gamma_1 n)!((1-\gamma_1)n)!} \left[\gamma_2{}^{\gamma_1}(1-\gamma_2)^{1-\gamma_1}\right]^n \left\{1 + \frac{1-\gamma_2}{\gamma_2}\frac{\gamma_1}{1-\gamma_1+1/n}\right\}$$

with a symmetric expression for q_3. Constructing the ratio, cancelling common terms and taking limits yields, for $\gamma_2 \neq 1/2$:

$$\lim_{n\to\infty} \log \frac{q_2}{q_3} = \lim_{n\to\infty} n(2\gamma_1 - 1)\log\left(\frac{\gamma_2}{1-\gamma_2}\right) = \left\{ \begin{array}{ll} +\infty & \gamma_2 > 1/2 \\ -\infty & \gamma_2 < 1/2 \end{array} \right\} \tag{1.5}$$

It follows that in the Cox-Palfrey case — unlike the encompassing model presented here — the incentive to vote strategically is unbounded in large constituencies.

1.6 Comparative Statics

The Cox-Palfrey model predicts complete strategic switching, and hence a bipartite outcome. Since strategic voting is always complete, it cannot change in response to the parameters. The only comparative static is discrete — changing the identity of the two leading contenders results in a different bipartite outcome.

Comparative statics are available for the model considered here. This section conducts such comparisons with the aid of the Beta prior described in Section 1.2.

1.6.1 Opinion Polls and Information Precision

Section 1.2 introduced the Beta prior as a convenient conjugate for the binomial distribution. In that section, the precision of the focal voter's beliefs carried the index m. A microfoundation is available, however. Begin with a Beta prior over p:

$$f(p) = \frac{\Gamma(\beta_2 + \beta_3)}{\Gamma(\beta_2)\Gamma(\beta_3)} p^{\beta_2 - 1}(1-p)^{\beta_3 - 1}$$

Suppose that the focal voter now observes an opinion poll of size m. This samples voters satisfying $u_1 < \min\{u_2, u_3\}$. Denote the realisation of this random sample by y_2 and $y_3 = m - y_2$. A standard application of the beta-binomial yields:

$$f(p\,|\,y) = \frac{\Gamma(\beta_2 + \beta_3 + m)}{\Gamma(\beta_2 + y_2)\Gamma(\beta_3 + m - y_2)} p^{y_2 + \beta_2 - 1}(1-p)^{m - y_2 + \beta_3 - 1}$$

Suppose that the share of candidate support in the opinion poll is γ_2, so that $y_2 = \gamma_2 m$. Assume without loss that $\gamma_2 > 1/2$, so that candidate 2 is expected to be the closest challenger. Specialise to a uniform prior, where $\beta_2 = \beta_3 = 1$. The posterior density becomes:

$$f(p\,|\,y) = \frac{\Gamma(m+2)}{\Gamma(\gamma_2 m + 1)\Gamma((1-\gamma_2)m + 1))} p^{\gamma_2 m}(1-p)^{(1-\gamma_2)m} \tag{1.6}$$

Compare Equation (1.6) to Equation (1.1, p. 6). The precision index m can thus be viewed as the size of an opinion poll observed by the focal voter; the modification in Equation (1.6) is due to the inclusion of the flat prior over p. For a large electorate ($n \to \infty$) the pivotal log-likelihood ratio becomes:

$$\lambda = \log \frac{f(\gamma_1)}{f(1-\gamma_1)} = (2\gamma_2 - 1)m \log\left(\frac{\gamma_1}{1-\gamma_1}\right) = (\gamma_2 - \gamma_3)m \log\left(\frac{\gamma_1}{1-\gamma_1}\right) \tag{1.7}$$

where $\gamma_3 = 1 - \gamma_2$. Inspecting Equation (1.7) yields the following proposition.

Proposition 2. *Strategic voting incentives are increasing in (a) the strength of candidate 1, (b) the perceived gap between candidates 2 and 3 and (c) the precision of information m.*

The comparative statics obtained here are mirrored by those obtained in the fully game-theoretic models of Parts 2 and 3. Indeed, as Section 2.1 notes, they challenge established intuition to some extent. Notice that in a "close" election, there is little incentive for strategic voting. Allowing $\gamma_1 \downarrow 1/2$ and $\gamma_2 - \gamma_3 \downarrow 0$ results in $\lambda \downarrow 0$. Perhaps the most important conclusion, however, is the insight offered into the rôle of information precision. The pivotal log-likelihood ration λ becomes unboundedly large as m increases. This effect is illustrated in Figure 1.5. More formally, taking limits:

$$\lim_{m\to\infty} \lambda = \lim_{m\to\infty} m(\gamma_2 - \gamma_3)\log\left(\frac{\gamma_1}{1-\gamma_1}\right) = +\infty \quad \text{for} \quad \min\{\gamma_1, \gamma_2\} > 1/2 \tag{1.8}$$

Notice the close correspondence between Equation (1.8) and the Cox-Palfrey limiting case in Equation (1.5). There are two key differences, however. First, the rôles of γ_1 and γ_2 are interchanged. Secondly — and more importantly — the Cox-Palfrey framework suggests that an unbounded constituency size is the key ingredient for a strict Duvergerian conclusion. Equation (1.8) demonstrates that this is not the case. Rather, it is an unbounded opinion poll size that is required.

Of course, as Figure 1.5 clearly illustrates, a relatively modest opinion poll is sufficient to obtain

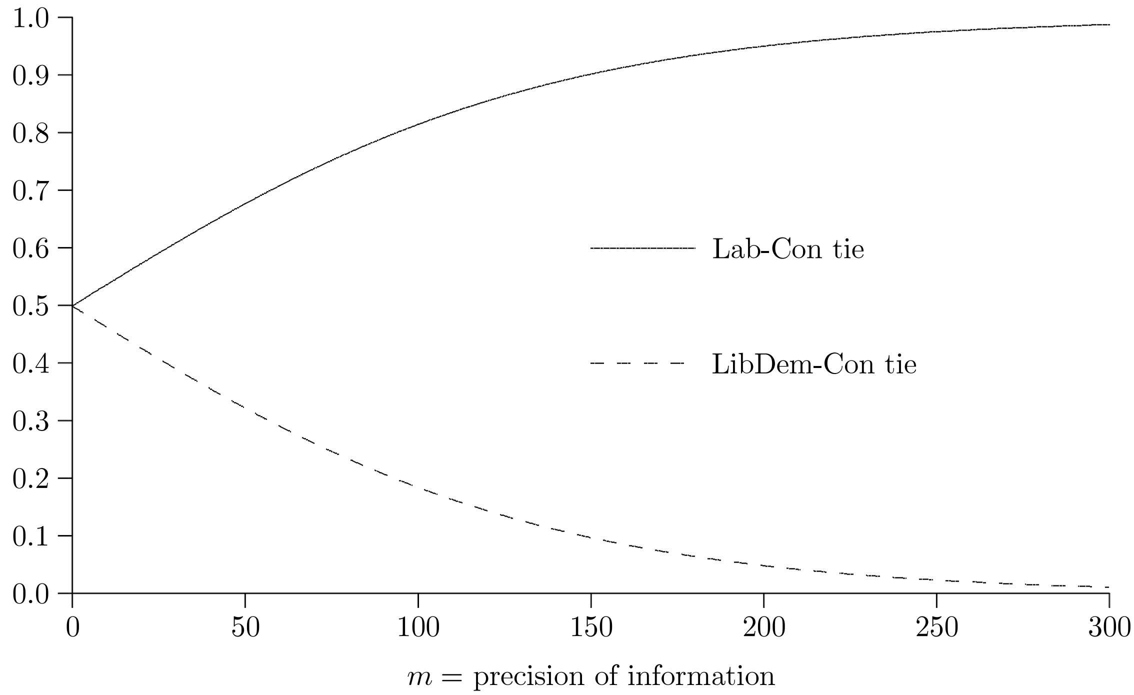


Figure 1.5: Conditional Pivotal Probabilities for $n \to \infty$ and Varying m

a near-Duvergerian outcome. This suggests that, in a decision-theoretic environment, one might expect near-complete strategic voting. Such a conclusion is in error. Part 3 argues that voting takes place at a *constituency* level, whereas high-precision public information sources (such as opinion polls) are typically seen at a *national* level. It is reasonable to expect the precision of information on constituency standings to be rather lower. In this case, the decision-theoretic model leads to a non-Duvergerian conclusion.

1.7 Conclusion

This first step argues that the key assumption driving the strict Duvergerian results of earlier formal voting models is the assumed independence of voter preferences. In a decision-theoretic framework, a non-Duvergerian outcome emerges. Perhaps most importantly, *constituency* uncertainty rather than *idiosyncratic* uncertainty is the key determinant of a strategic voter's decision.

Issues remain. The analysis presented here is decision-theoretic — voters do not take the strategic decisions of others into account. A game-theoretic perspective is required. Indeed, one might expect strategic voting to be self-reinforcing, and hence the possibility of fully Duvergerian game theoretic conclusion remains open. Furthermore, the constituency-level uncertainty lacks a microfoundation. Finally, the information sources available to the electorate have yet to be fully explored.

These issues are addressed in the following partss. Part 2 provides a foundation for constituency uncertainty, by breaking voter preferences into common and idiosyncratic components. Moreover, direct modelling of preferences allows a full game-theoretic analysis. The insights carry through — with sufficient precision a Duvergerian conclusion emerges. Part 3 generalises the model further with the introduction of privately observed signals. It argues that, at a constituency level, this is a more plausible framework. Surprisingly, it is shown that strategic voting exhibits negative feedback, and is self-attenuating. Non-Duvergerian conclusions follow.

1.8 Omitted Results

1.8.1 Pivotal Probabilities with Uncertain p (Section 1.4.2)

This section outlines the procedure used to generate Example 2 and Figure 1.3. The prior distribution over p is the Beta distribution, with parameters $\beta_2 = m\gamma_2$ and $\beta_3 = m(1-\gamma_2)$, where for illustration $m = 40$. A standard application of the distribution shows that:

$$\begin{aligned} \Pr[x_2 = \gamma_1 n] &= \frac{\Gamma(\beta_2+\beta_3)}{\Gamma(\beta_2)\Gamma(\beta_3)} \frac{n!}{(\gamma_1 n)!((1-\gamma_1)n)!} \int_0^1 p^{\gamma_1 n+\beta_2-1}(1-p)^{(1-\gamma_1)n+\beta_3-1}\, dp \\ &= \frac{\Gamma(\beta_2+\beta_3)}{\Gamma(\beta_2)\Gamma(\beta_3)} \frac{\Gamma(n+1)}{\Gamma(\gamma_1 n+1)\Gamma((1-\gamma_1)n+1)} \frac{\Gamma(\gamma_1 n+\beta_2)\Gamma((1-\gamma_1)n+\beta_3)}{\Gamma(\beta_2+\beta_3+n)} \end{aligned}$$

It follows that:

$$\frac{\Pr[x_2 = \gamma_1 n]}{\Pr[x_3 = \gamma_1 n]} = \frac{\Gamma(\gamma_1 n+\beta_2)\Gamma((1-\gamma_1)n+\beta_3)}{\Gamma(\gamma_1 n+\beta_3)\Gamma((1-\gamma_1)n+\beta_2)}$$

Figure 2 plots this ratio for varying n. This is not the exact ratio q_2/q_3, since it omits near ties. Inclusion of exact ties in this illustration has a neglible effect.

1.8.2 Omitted Proofs from Section 1.5

Proof of Lemma 2: Recall from Section 1.3 that optimal voting behaviour is contingent upon the relative probabilities of a tie for the lead. Such a tie requires $\gamma_1 n$ voters to coordinate on a candidate $j \in \{2, 3\}$. In addition, the event of an "almost tie" is also of interest. Thus:

$$q_j = \Pr[\gamma_1 n - 1 \leq x_j \leq \gamma_1 n]$$

As in Section 1.2, denote by p the (conditionally independent) probability that a voter supports candidate $j = 2$. Thus, conditional on p, the vote total x_2 is binomially distributed with parameters p and n:

$$\Pr[x_2 = \gamma_1 n \,|\, p] = \frac{n!}{(\gamma_1 n)!((1-\gamma_1)n)!} \left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n$$

From the perspective of the focal voter, p is unknown. It follows that:

$$\Pr[x_2 = \gamma_1 n] = \frac{n!}{(\gamma_1 n)!((1-\gamma_1)n)!} \int_0^1 \left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n f(p)dp$$

Furthermore:

$$\Pr[x_2 = \gamma_1 n - 1] = \frac{n!}{(\gamma_1 n)!((1-\gamma_1)n)!} \frac{\gamma_1 n}{n(1-\gamma_1)+1} \int_0^1 \left[p^{\gamma_1}(1-p)^{1-\gamma_1}\right]^n \frac{1-p}{p} f(p)dp$$

Combining these yields the appropriate expression. It follows that a symmetric expression holds for q_3. $\square$

Proof of Lemma 3: Employ the notation used in the proof of Lemma 4. Hence, for $j = 2$:

$$\frac{\Pr[x_2 = \gamma_1 n - 1]}{\Pr[x_2 = \gamma_1 n]} = \frac{\gamma_1}{1-\gamma_1+1/n} \frac{\int_0^1 G(p)^n \frac{1-p}{p} f(p)dp}{\int_0^1 G(p)^n f(p)dp}$$

The limit of the first term is $\gamma_1/(1-\gamma_1)$. An upper bound for the second term is:

$$\frac{f_{H,\epsilon}(\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon} G(p)^n \frac{1-p}{p} dp + f_H\left[\int_0^{\gamma_1-\epsilon} G(p)^n \frac{1-p}{p} dp + \int_{\gamma_1+\epsilon}^1 G(p)^n \frac{1-p}{p} dp\right]}{f_{L,\epsilon}(\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon} G(p)^n dp}$$

Taking the first term in this expression:

$$\frac{f_{H,\epsilon}(\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon} G(p)^n \frac{1-p}{p} f(p)dp}{f_{L,\epsilon}(\gamma_1)\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon} G(p)^n dp} \le \frac{1-\gamma_1+\epsilon}{\gamma_1-\epsilon} \frac{f_{H,\epsilon}(\gamma_1)}{f_{L,\epsilon}(\gamma_1)}$$

Next consider the second term:

$$\frac{\int_0^{\gamma_1-\epsilon} G(p)^n \frac{1-p}{p} dp}{\int_{\gamma_1-\epsilon}^{\gamma_1+\epsilon} G(p)^n dp} \le \frac{1}{\epsilon} \int_0^{\gamma_1-\epsilon} \left[\frac{G(p)}{G(\gamma_1-\epsilon)}\right]^n \frac{1-p}{p} dp \longrightarrow 0$$

Similarly this holds for the second term. It follows that:

$$\lim_{n\to\infty} \frac{\Pr[x_2 = \gamma_1 n - 1]}{\Pr[x_2 = \gamma_1 n]} \le \lim_{\epsilon\to\infty} \left\{\frac{\gamma_1}{1-\gamma_1} \frac{1-\gamma_1+\epsilon}{\gamma_1-\epsilon} \frac{f_{H,\epsilon}(\gamma_1)}{f_{L,\epsilon}(\gamma_1)}\right\} = 1$$

Once again, a symmetric procedure bounds the limit below, and the result obtains. $\square$

Part 2

Strategic Voting with Constituency Uncertainty

"In these matters the only certainty is that there is nothing certain."

Pliny the Elder

2.1 Introduction

2.1.1 The Rôle of Constituency Uncertainty

Farquharson (1969, pp. 57–60) notes that interest in strategic voting dates back to Pliny the Younger. The comment of Pliny the Elder recorded here is perhaps apt in advising the modeller in the pursuit of its analysis. Indeed, the critique of Part 1 suggests that uncertainty over constituency-wide support of candidates is the key missing ingredient in the Cox (1994) and Palfrey (1989) models of strategic voting.

The decision-theoretic analysis of Part 1 leads to a non-Duvergerian conclusion. Such an analysis is incomplete, as it omits the effect of strategic voting by others on an individual's voting decision. A game-theoretic perspective is required, necessitating a full specification of voter preferences. Furthermore, the source of constituency-uncertainty is unclear. A microfoundation for this uncertainty is desirable.

To address these issues, the simple model of a three candidate plurality election described in Part 1 is extended. Voter preferences are fully specified. The model has much in common with those of Cox (1994) and Fey (1997), employing a random utility framework. The key innovation is the presence of a common utility element to the payoffs of each voter. The value of this common element is uncertain. This, in turn, induces uncertainty over the distribution of preferences in the population, and is equivalent to uncertainty about the popular support of candidates. It also captures the notion that there are constituency-wide factors which affect the preferences of the whole electorate, as well as idiosyncratic factors that are voter-specific. The remainder of this section discusses the approach in more detail.

2.1.2 Constituency-Wide Common Effects

When will the independence of voting decisions fail? The preference specification of the new theory suggests a possible answer. Preferences are decomposed into *common* and *idiosyncratic* components. The common component affects all voters in a constituency. Idiosyncratic components are individual to a voter, and distributed independently across the electorate. This implies that, conditional on the common preference component, voting decisions are independent. This common component is, however, unknown to an observer — including the members of the electorate. It follows that unconditional preferences (and hence voting decisions) are correlated. The decision of one individual provides a signal of the common utility component, and hence the voting decisions of other individuals. Existing models (Cox 1994, Palfrey 1989, Fey 1997) are equivalent to this specification when the common component (and hence constituency-wide support) is commonly known.

Part 1 argues that it is *only* uncertainty over constituency-wide support that drives strategic incentives. Indeed, with the microfoundation of common utility elements, the analysis here shows that it is uncertainty over common effects that matters. Once again, individual idiosyncrasies are averaged out in a large electorate, and hence do not influence the likelihood ratios of tying pairs.

2.1.3 Self-Reinforcing Strategic Voting

Adopting a game-theoretic perspective, the possibility of a strict Duvergerian outcome remains open. Strategic switching from trailing candidates and toward the leading pair reduces the probability that trailing candidates are pivotal. This further increases the incentives of instrumentally rational voters to switch their support to the leading pair. Further switching reinforces this effect. Where will such a process lead? There are two possibilities. First, the process might continue until all support for trailing candidates is lost. Alternatively, the process may converge to an equilibrium in which the trailing candidates retain some attenuated support. Untrained intuition is insufficient to determine which of these situations obtain, and a formal analysis is required. Such analysis necessitates a focus on the information source on which voters base their decisions.

To explore this issue, voters are equipped with a public information source. All individuals commonly observe a *public* signal of the common component of preferences — a public signal of constituency-wide support for the candidates. This may be likened to the observation of an opinion poll, and in fact a detailed opinion poll provides the microfoundation. Naturally, voters have at least some *privately* observed information to assist them in their decisions. Absent communication with any specific outside source, a voter might engage in introspection, examining their own preferences to elicit a signal of the common component. In this part, such factors are ignored, and voters are assumed to discard their personal preferences when formulating their beliefs. This omission is rectified in Part 3, with somewhat surprising results.

With pure public information, where voters ignore any private signal, strategic voting is indeed self-reinforcing. In the three-candidate model considered here, the analysis identifies three equilibria. The two Duvergerian equilibria entail complete coordination on one of the challenging candidates, yielding a bipartite outcome. There is also a unique non-Duvergerian equilibrium. An equilibrium selection problem is posed.

Equilibrium selection is made according to a stability criterion. If the public signal is sufficiently noisy, then the non-Duvergerian equilibrium is uniquely stable. Beginning with an initial "straightforward voting" hypothesis and engaging in iterative best response, strategic voting exhibits positive feedback, but peters out as an interior equilibrium is attained. By contrast, when the public signal is sufficiently precise, strategic voting is explosive, and all support for the trailing candidate is lost as the best-response process converges to a Duvergerian equilibrium. Existing models correspond to this latter Duvergerian case — a Duvergerian equilibrium requires public information of sufficient precision. In both cases, strategic voting is uni-directional.

This selection device mimics the repeated election dynamic described by Fey (1997, pp. 142–146). Indeed, as voter beliefs are allowed to become unboundedly precise, the Fey model is obtained. Notice, however, that repeated elections are unnecessary to achieve the desired selection. Beginning with an initial hypothesis of straightforward voting, the deliberation dynamic alone is sufficient to attain an equilibrium outcome.

In a Duvergerian election, the analysis of this part lends qualified support to strict bipartism. This requires, however, sufficient precision of voter beliefs and hence sufficient detail in a public signal. As will become clear in Part 3, this Duvergerian conclusion is quickly overturned following the introduction of private information.

2.1.4 Comparative Statics

The comparative statics of the game-theoretic model mirror those of the decision-theoretic model of Part 1. In particular, there is a key rôle for the precision of information. If information is sufficiently precise, then a fully Duvergerian outcome obtains. In the non-Duvergerian case, strategic voting is increasing in the precision of the public signal.

As in Part 1, the impact of the expected strengths of the various candidates may be ascertained. As one might expect, the incentive to vote strategically is increasing in the expected gap between challenging candidates. Informal intuition has previously suggested that strategic voting increases as the lead of a leading candidate weakens. Cain (1978, p. 644) provides a classic example in his analysis of strategic voting in Britain:

> "we would expect the level of third-party support to be lower in competitive (i.e. closely contested) than in noncompetitive constituencies (i.e. one party dominant), since the pressure to defect and case a meaningful vote will be greater in constituencies with close races."

This intuition fails. In fact, the incentive to vote strategically increases with the strength of a leading candidate. Intuitively, greater coordination is required to defeat this leading candidate, and this increases strategic incentives. Cain's argument appears to be based on the *absolute* probability of influencing the election outcome. But as the analyses of both the new theory and its predecessors demonstrate, it is only the *relative* likelihood of tied outcomes that determines the optimal voting decision.

In this fully-specified model, an additional parameter is available — the heterogeneity of voter preferences. The initial (decision-theoretic) incentive to vote strategically is unaffected by the heterogeneity of voters. The *amount* of strategic voting, however, increases as the heterogeneity of the population is reduced. Intuitively, if there is greater commonality of opinion, then only a limited incentive is required to encourage a large number of individuals to switch their support.

2.1.5 Guide

The argument continues in the following sections. Section 2.2 extends the model of Section 1.2 to include a full specification of voter preferences, as well as the public signal observed by voters. Section 2.3 investigates the optimal behaviour of voters, and calculates the realised pivotal log-likelihood ratio given such behaviour. This allows an immediate decision-theoretic analysis. Equilibrium considerations are introduced in Section 2.4, where the perceived and realised pivotal log-likelihood ratios are equated. Possible equilibria are classified, and equilibrium selection is made.

2.2 Model

2.2.1 The Election

The specification of the election is identical to that of part 1. The reader is referred to Section 1.2 of that part for full details. There are three candidates $j \in \{1, 2, 3\}$ in a single-seat district selection, operating under the plurality rule. Realised vote totals will be denoted $\{x_j\}$, where $\sum x_j = (1+\gamma_1)n+1$. The vote total for candidate 1 is fixed at $x_1 = \gamma_1 n$, and remaining voters prefer both of candidates 2 and 3 to candidate 1. Assumption 1 is upheld ($\gamma_1 > 1/2$), ensuring that any tie for the lead involves candidate 1. All references to a "voter" and the "electorate" will refer to the $n+1$ remaining voters.

2.2.2 Preference Structure

A full specification of voter prefereces is unnecessary for a decision theoretic analysis. Given the beliefs of a focal voter (represented in Part 1 by the density $f(p)$), the pivotal log-likelihood ratio λ may be calculated. The strategic incentive may then be investigated relative to the fixed preferences of any particular focal voter.

In a game-theoretic environment, additional structure is required. A focal voter must consider the probability of a strategic switch by other voters. This probability depends on both the perceived log-likelihood ratio and the distribution of voter preferences. To address this need, a random utility model is adopted. The utilities include common effects, which individual voters do not observe. Once again, the restriction will be to the n voters satisfying $u_{i1} < \min\{u_{i2}, u_{i3}\}$.

Notation 1. *Voter i has von Neumann-Morgenstern utility u_{ij} for party j. This satisfies:*

$$\log \frac{u_{i2} - u_{i1}}{u_{i3} - u_{i1}} = \eta + \epsilon_i \tag{2.1}$$

Refer to η and ϵ_i as the common *and* idiosyncratic *components of utility.*

Notice that the left hand side of Equation (2.1) gives a scale-free measure of the preference for candidate 2 relative to candidate 3, each measured relative to the benchmark of candidate 1. Section 1.3 of Part 1 demonstrates that this measure is sufficient for the optimal voting decision. Without loss of generality, it is assumed that $\eta > 0$, so the actual constituency-wide support for candidate 2 is greater.

Consider next the decomposition on the right hand side of Equation (2.1). An easy interpretation is that η represents constituency-wide factors affecting all voters. By contrast, ϵ_i represents the idiosyncratic preference of voter i. More generally, η is the expectation of $\log(u_{i2} - u_{i1}) - \log(u_{i3} - u_{i1})$ conditional on all district-level information, generating the residual component ϵ_i. This residual will be independent across voters. In addition, a parametric specification is employed. This is not critical to the argument, but rather allows an easy microfoundation for the public signal described in Section 2.2.3.

Assumption 3. *The idiosyncratic component ϵ_i is distributed normally: $\epsilon_i \sim N(0, \xi^2)$.*

If the common component η is fixed and commonly known, then voters' preferences are independent. Such an assumption yields the models of Cox (1994) and Fey (1997). In particular, it implies certain knowledge of constituency-wide preferences and hence the election outcome. To see this, notice that the probability that a voter's first preference is for party 2, conditional on η, is $\gamma_2 = \Phi(\eta/\xi)$, where Φ is the cumulative distribution function of the standard normal. Defining $\gamma_3 = 1 - \gamma_2$, it follows that, with truthful voting, $\mathrm{E}[x_j] = \gamma_j n$. Indeed, fixing the constituency-wide truthful support levels for each candidate, the common utility component satisfies $\eta = \xi \Phi^{-1}(\gamma_2)$.

If η is fixed and known, then the expected (truthful) support for all candidates is commonly known throughout the constituency. In a large constituency, enforcement of the Law of Large Numbers reveals that the outcome of any election with truthful voting is commonly known by all constituents — a restrictive assumption indeed. The key innovation of the new theory is the introduction of uncertainty over η, and hence uncertainty over the support for each candidate. This provides a microfoundation for the uncertainty over voting decisions represented in Part 1 by the density $f(p)$.

2.2.3 Public Signals

Voters begin with a diffuse prior over η — they have no knowledge of average candidate support. Information on candidate support may then be gleaned from a variety of sources. This part specialises to the case of *public signals*. One important source of public information is the publication of opinion polls and similar surveys. Formally, this is modelled as the receipt of a public signal, equal to the true value of η plus noise.

Assumption 4. *Voters commonly observe a* public signal $\mu \sim N(\eta, \sigma^2)$.

Following observation of this signal, and prior to the receipt of any private information, voters update to a common public posterior $\eta \sim N(\mu, \sigma^2)$. A microfoundation underpins Assumption 4. Suppose that the preferences of m_μ randomly chosen invididuals are publicly observed. Indexing these individuals by $k \in \{1, 2, \ldots, m_\mu\}$:

$$\mu = \frac{1}{m_\mu} \sum_{k=1}^{m_\mu} \log \frac{u_{k2} - u_{k1}}{u_{k3} - u_{k1}} \sim N\left(\eta, \frac{\xi^2}{m_\mu}\right) \tag{2.2}$$

Thus a public signal with variance $\sigma^2 = \xi^2/m_\mu$ is equivalent to the public observation of m_μ individuals. The generation of Equation (2.2) employs the distributional specification of Assumption 3. For large m_μ, however, the Central Limit Theorem suggests the normal as an appropriate specification for the distribution of μ.

Viewed as an opinion poll, Assumption 4 provides a natural framework. In particular, the widespread publication of opinion polls is common during an election. Implicit to this interpretation, however, is the assumption that sampled individuals reveal their true preferences to the pollster. This issue is neglected here, and for good reason. As Part 3 argues, it is information on the *district* level candidate support that is of interest. Opinion polls are rarely conducted at the constituency level in countries such as the United Kingdom, and hence it is perhaps privately observed information that is of more importance. It is argued that, in this case, any signal is elicited from the observation of society over a longer time period, with reduced likelihood of strategic manipulation of viewpoints on the part of respondents. Of course, this justification immediately calls into question the relevance of a model based entirely on public information signals. The aim of this part is to provide a public-information benchmark to the full model of Part 3. The reader is invited to view the present formulation in this light, and reserve judgement until the conclusion of the argument in the following part.

Of course, even in the absence of any explicit private signals, a private information source is implicitly available. As Section 2.1 noted, an individual's own payoffs provide a private signal of the common utility element. Any rational voter will employ all information available to them in order to reach a decision. Pursuing the benchmark justification still further, however, this element is neglected here. Indeed, voters are explicitly assumed to form beliefs over the common component η using only the public signal. It follows that all voters share a common public posterior belief:

$$\eta \sim N(\mu, \sigma_2)$$

This posterior may be obtained by beginning with a common public prior, Bayesian updating, and allowing the precision of the prior to vanish.

2.3 Voting Behaviour

2.3.1 Optimal Voting

As Section 1.3 observes, a vote can only influence the outcome of the election if it is *pivotal*. A pivotal situation arises with a tie or near-tie for the lead. Lemma 1 deduced the optimal voting rule for a voter, conditional on the perceived probabilities q_2 and q_3 of either a 1-2 tie or a 1-3 tie. Indeed, recalling Definition 2 of the pivotal log-likelihood ratio $\lambda = \log(q_2/q_3)$, this optimal voting rule satisfies:

$$\text{Vote } 2 \;\; \Leftrightarrow \;\; \log \frac{u_{i2} - u_{i1}}{u_{i3} - u_{i1}} + \lambda \geq 0$$

The analysis focuses on symmetric strategy profiles. This is without loss of generality, since voters will act optimally, and their beliefs will be based entirely on the public signal μ. Conditioning on the perceived log-likelihood ratio λ, which is common to all voters, and the common utility component, an individual supports candidate 2 whenever:

$$\eta + \epsilon_i + \lambda \geq 0 \tag{2.3}$$

This yields the following simple lemma.

Lemma 5. *Conditional on η and λ, a voter supports candidate 2 with probability:*

$$p = \Pr\left[\log \frac{u_{i2} - u_{i1}}{u_{i3} - u_{i1}} + \lambda \geq 0 \,|\, \eta, \lambda\right] = \Phi\left(\frac{\eta + \lambda}{\xi}\right)$$

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

Proof. Conditional on η and λ, $\epsilon_i \sim N(0, \xi^2)$. Re-arrange Eq. (2.3) to obtain the result. □

This implicitly assumes that $\lambda \neq \pm\infty$. These cases will be considered in Section 2.4.

Suppose that all voters share a common perception λ. What will be the *realised* pivotal log-likelihood ratio $\hat{\lambda}$? A symmetric strategy profile yields an identical perception λ of the log-likelihood ratio of pivotal probabilities for each voter. Fixing this perception, and hence the strategy profile, yields a realised log-likelihood ratio $\hat{\lambda}$. This depends in turn on the probability that a voter supports candidate 2 (p), and uncertainty over this probability. Represent this uncertainty by the density $f(p)$, and assume for now that this is bounded above. Allowing the constituency size n to grow large, Proposition 1 from Part 1 reveals:

Proposition 3. *Allowing the electorate size to grow large:*

$$\lim_{n\to\infty} \hat{\lambda} = \log \frac{f(\gamma_1)}{f(1-\gamma_1)}$$

This requires evaluation of the density $f(p)$. Uncertainty over p is driven by uncertainty over η, and yields a convenient formula for $\hat{\lambda}$, as the next proposition reveals.

Proposition 4. *With pure public information and $\lambda \neq \pm\infty$ the realisation $\hat{\lambda}$ satisfies:*

$$\hat{\lambda}(\lambda) = \frac{2\xi\Phi^{-1}(\gamma_1)(\lambda+\mu)}{\sigma^2} \tag{2.4}$$

Proof. Conditioning on η and λ, Lemma 5 reveals that:

$$p = \Pr\left[\log \frac{u_{i2}-u_{i1}}{u_{i3}-u_{i1}} + \lambda \geq 0 \,|\, \eta\right] = \Phi\left(\frac{\eta+\lambda}{\xi}\right)$$

Hence:

$$F(p) = \Pr\left[\Phi\left(\frac{\eta+\lambda}{\xi}\right) \leq p\right] = \Pr[\eta \leq \xi\Phi^{-1}(p) - \lambda] = \Phi\left(\frac{\xi\Phi^{-1}(p)-\lambda-\mu}{\sigma}\right)$$

Differentiate this expression with respect to p to obtain $f(p)$:

$$f(p) = \frac{\xi}{\sigma\phi(\Phi^{-1}(p))}\phi\left(\frac{\xi\Phi^{-1}(p)-\lambda-\mu}{\sigma}\right)$$

Notice that this density is bounded above, as required. Next, evaluate this density at γ_1 and $1-\gamma_1$. Exploiting the symmetry of the normal distribution, note that $\Phi^{-1}(1-\gamma_1) = -\Phi^{-1}(\gamma_1)$. Furthermore, $\phi(z) = \phi(-z)$. It follows that $\phi(\Phi^{-1}(1-\gamma_1)) = \phi(\Phi^{-1}(\gamma_1))$. Taking ratios:

$$\frac{f(\gamma_1)}{f(1-\gamma_1)} = \exp\left(\frac{(\xi\Phi^{-1}(\gamma_1)+\lambda+\mu)^2 - (\xi\Phi^{-1}(\gamma_1)-\lambda-\mu)^2}{2\sigma^2}\right)$$

Simplifying the numerator in the exponent and taking logarithms yields the desired result. □

2.3.2 Decision-Theoretic Voting

This proposition immediately allows a decision-theoretic conclusion. McKelvey and Ordeshook (1972) consider the optimal response of a strategic voter, assuming that all other voters act straightforwardly. A straightforward voter chooses to vote for their preferred candidate. Naturally, this is equivalent to voting optimally with a perceived pivotal log-likelihood ratio of $\lambda = 0$.

Employing Proposition 4, a realised log-likelihood ratio λ_{DT} obtains:

$$\lambda_{DT} = \hat{\lambda}(0) = \frac{2\xi\Phi^{-1}(\gamma_1)\mu}{\sigma^2}$$

This provides a measure of the incentive for a decision-theoretic strategic voter to switch support away from a preferred candidate, given the assumption that all other voters act straightforwardly. This measure requires an appropriate normalisation. First, recall that the microfoundation for public signals specified that $\sigma^2 = \xi^2/m_\mu$, where m_μ is the size of a publicly observed survey of voter preferences. This yields:

$$\lambda_{DT} = \frac{2m_\mu\Phi^{-1}(\gamma_1)\mu}{\xi}$$

It is immediate that the incentive to vote tactically is increasing in the precision of public information (m_μ), the strength of the first candidate (γ_1) and the perceived distance between the second and third candidates (μ). It is tempting to conclude that the incentive is also decreasing in idiosyncratic variability (ξ). This conclusion requires more thought, however. Varying ξ also changes the expected (truthful) support of candidate 2, since $\gamma_2 = \Phi(\eta/\xi)$. A more appropriate comparative static is to fix γ_2, and allow η to vary with ξ. In this case:

$$\mathrm{E}[\lambda_{DT}] = \frac{2m_\mu\Phi^{-1}(\gamma_1)\eta}{\xi} = 2m_\mu\Phi^{-1}(\gamma_1)\Phi^{-1}(\gamma_2) \tag{2.5}$$

Similarly, the variance of λ_{DT} is invariant to ξ. It follows that ξ has no direct effect on the incentive to vote strategically. More interestingly, however, the *probability* that a voter acts strategically is decreasing in ξ. To see this, note that an individual acts strategically whenever $\eta + \epsilon_i + \lambda_{DT} \geq 0 \geq \eta + \epsilon_i$. In particular, the "average" individual ($\epsilon_i = 0$) always votes strategically. For small ξ, many more inviduals satisfy this inequality, for fixed η and λ_{DT}. In fact, conditioning on an accurate signal realisation of $\mu = \eta$, the probability of a strategic vote is:

$$\Phi\left(\Phi^{-1}(\gamma_2)\left(1 + \frac{2m_\mu\Phi^{-1}(\gamma_1)}{\xi}\right)\right) - \gamma_2$$

This is clearly decreasing in ξ.

Additional observations are pertinent. First, notice that in a "close" election, the incentive for (and amount of) strategic voting is small. An election becomes more marginal as γ_1 decreases to $1/2$. It follows that $\Phi^{-1}(\gamma_1) \downarrow 0$, and the incentive to vote strategically disappears. Secondly, as candidates 2 and 3 become closer, $\eta \downarrow 0$, and the expected incentive again disappears, with the probability of a strategic vote. The model predicts less strategic voting in a close election. Third, notice that the incentive (and hence amount) of strategic voting is finite. Uncertainty over η leads to a bounded likelihood ratio of pivotal probabilities. The decision-theoretic outcome is thus distinctly non-Duvergerian in nature. The final observation is to note that the decision-theoretic outcome tends to the Duvergerian case as the precision of public information increases. More formally:

$$\mu > 0 \;\Rightarrow\; \lim_{m_\mu \to \infty} = +\infty$$

It follows that the Duvergerian conclusions of previous decision-theoretic analysis hinge on exact precision of public information — equivalent to complete and common knowledge of constituency-wide preferences. These results mimic those of Part 1, with the addition of fully specified voter preferences.

Of course, full insight requires a game-theoretic analysis. In the next section, the argument considers the properties of equilibria in this model.

2.4 Equilibrium

2.4.1 Equilibrium Voting

The decision-theoretic analysis considered the nature of optimal voting, assuming that other voters act truthfully. A rational voter will anticipate strategic behaviour by others, leading to a revision of the perceived log-likelihood λ. Linking the optimal voting rule with the standard consistency requirement yields a Bayesian Nash equilibrium. The previous section was concerned with finite λ, and first attention must turn to the possibility of extreme Bayesian Nash equilibria.

Proposition 5. *There exist two Duvergerian Bayesian Nash equilibria, with $x_2 = n + 1$ and $x_3 = n + 1$ respectively, resulting in a certain win by either candidate 2 or candidate 3.*

Proof. Suppose that all $n + 1$ voters satisfying $u_{i1} < \min\{u_{i2}, u_{i3}\}$ vote for candidate 2. The result is a certain win by candidate 2. There is exactly zero probability of a tie, so that $q_2 = q_3 = 0$. Voters receive the same expected payoff from a vote for either party $j \in \{2, 3\}$, and there is no incentive to deviate. An identical argument establishes that $x_3 = n + 1$ is also an equilibrium. □

Under Duvergerian voting, the log-likelihood ratio λ is undefined, since there is exactly zero probability of a tie for the lead. The key question of interest is the existence of non-Duvergerian Bayesian Nash equilibria. In this case, with finite λ, the consistency of voter perceptions and the realised distribution over voter outcomes may be imposed.

Proposition 6. *There is a unique symmetric non-Duvergerian equilibrium λ^* satisfying:*

$$\lambda^* = \frac{2\xi\Phi^{-1}(\gamma_1)\mu}{\sigma^2 - 2\xi\Phi^{-1}(\gamma_1)} \tag{2.6}$$

Proof. Set $\lambda^* = \hat{\lambda}(\lambda^*)$ in Equation (2.4) and solve. □

In common with the decision theoretic case, this non-Duvergerian equilibrium entails no strategic voting whenever the lead of candidate 1 is low ($\gamma_1 \downarrow 1/2$) and when the expected support for candidates 2 and 3 is the same ($\mu = 0$). Two issues remain, however. First, an equilibrium selection problem arises. Second, the direction of strategic voting is unclear. In the decision-theoretic case, $\mu > 0$ led to $\lambda_{DT} > 0$ and hence strategic voting away from candidate 3, as one might expect. The same conclusion is not immediate here. The denominator of Equation (2.6) may be either positive or negative, and the direction of switching is ambiguous.

2.4.2 Equilibrium Selection

The resolution of both of these problems is aided by the following definition.

Definition 3. *An election is* Duvergerian *if* $\sigma^2 < 2\xi\Phi^{-1}(\gamma_1)$, non-Duvergerian *if* $\sigma^2 > 2\xi\Phi^{-1}(\gamma_1)$.

The issue of equilibrium selection is first addressed. Suppose that voters have an initial hypothesis for decision rule employed by the electorate, yielding log-likelihood ratio λ_0 — for example, truthful voting corresponds to $\lambda_0 = 0$. A (decision-theoretic) best-response to this yields a realised log-likelihood ratio of $\lambda_1 = \hat{\lambda}(\lambda_0)$. A sophisticated voter may anticipate this response throughout the electorate. If all voters anticipate similarly, the realised log likelihood ratio becomes $\lambda_2 = \hat{\lambda}(\lambda_1)$. Continuing iteratively, a best-response process is obtained.

Definition 4. *Define the* iterative best response process *by* $\lambda_t = \hat{\lambda}(\lambda_{t-1})$.

Previous authors (including Fey 1997) specified a process of repeated elections. Voters observe the outcome of an election, and adopt a best response to the behaviour represented by the election outcome. In a second election, more strategic voting is observed. This process is reinforcing, leading to Duvergerian equilibria. With instrumentally rational agents, repeated elections are unnecessary for this process. A rational voter will anticipate the strategic actions of others, and further anticipate their reaction. Naturally, the limit of such a sophisticated reasoning process requires the specification of a starting point. An initial hypothesis of truthful voting provides a reasonable and intuitive focal point for such an initial condition.

Proposition 7. *In a non-Duvergerian election, the iterative best response process converges to the unique non-Duvergerian equilibrium from any initial λ_0. In a Duvergerian election, the iterative best response converges to a Duvergerian equilibrium, unless $\lambda_0 = \lambda^*$. If $\lambda_0 > \lambda^*$, it converges to the Duvergerian equilibrium in which candidate 3 receives no support. If $\lambda_0 < \lambda^*$, it converges to zero support for candidate 2.*

Proposition 7 is best understood with reference to Figures 2.1 and 2.2. Consider first the non-Duvergerian case, and suppose that the initial hypothesis is that voters act truthfully. This is

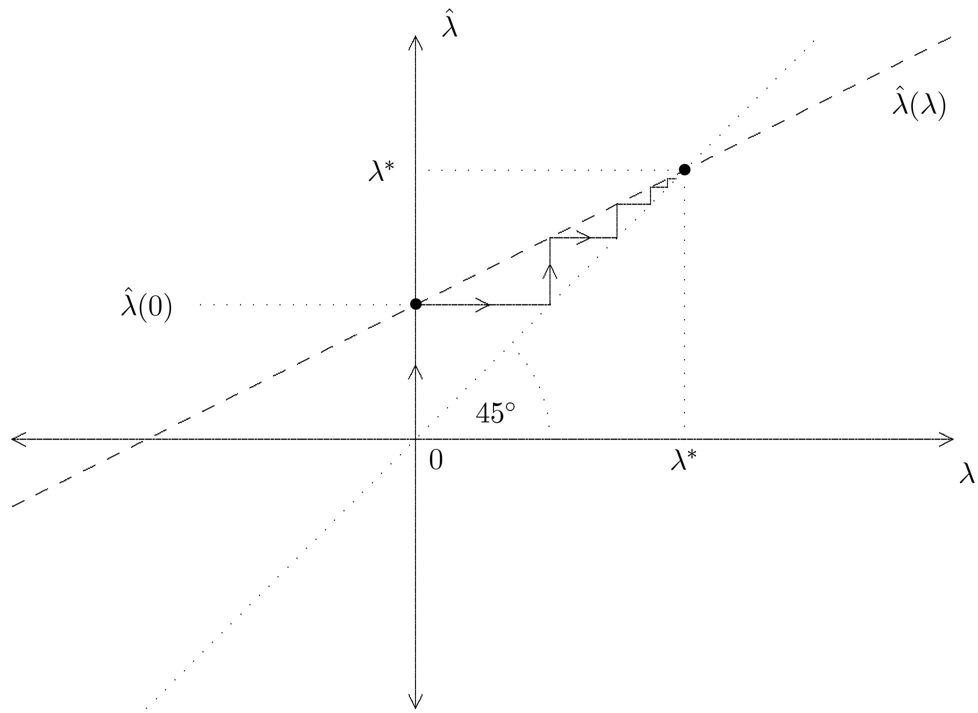


Figure 2.1: Convergence to a Stable Non-Duvergerian Equilibrium

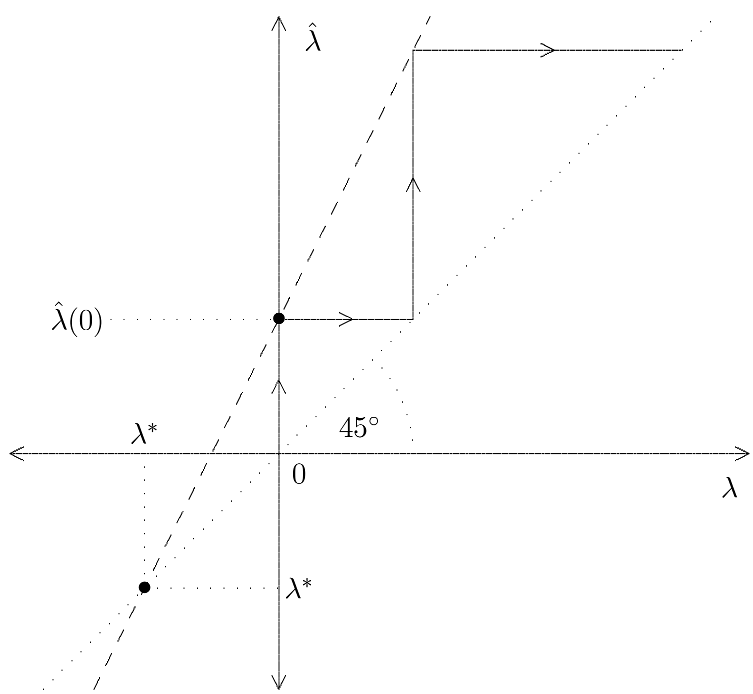


Figure 2.2: Divergence from an Unstable Non-Duvergerian Equilibrium

equivalent to optimal voting with $\lambda_0 = 0$, and a best response to this is $\lambda_{DT} = \hat{\lambda}(0)$. As Figure 2.1 illustrates, the best-response process continues, and converges to the non-Duvergerian equilibrium λ^*. That the initial hypothesis is truthful voting is unnecessary in a non-Duvergerian election — as the figure indicates and Proposition 7 confirms, any initial non-Duvergerian hypothesis leads to the non-Duvergerian equilibrium. Figure 2.2 illustrates the Duvergerian case. Notice that $\lambda^* < 0$ when $\mu > 0$. Beginning at $\lambda_0 = 0$, the reasoning process leads explosively away, and support for the third candidate collapses. Proposition 7 is employed as an equilibrium selection device, supporting the non-Duvergerian equilibrium if and only if the non-Duvergerian criterion is satisfied.

A restatement of the model in terms of the public survey microfoundation is instructive. Again noting that $\sigma^2 = \xi^2/m_\mu$, the criterion for a non-Duvergerian election (and hence equilibrium outcome) becomes:

$$\frac{1}{m_\mu} > \frac{2\Phi^{-1}(\gamma_1)}{\xi}$$

It follows that an election is Duvergerian if and only if the public signal of candidate support is sufficiently precise, relative to a measure of the support for candidate 1 (γ_1) and the heterogeneity of voter preferences (ξ). Notice that a Duvergerian outcome obtains if candidate 1 is sufficiently strong (large γ_1). In addition, increased heterogeneity of preferences (large ξ) yields a non-Duvergerian classification. The non-Duvergerian equilibrium log-likelihood ratio becomes:

$$\lambda^* = \frac{2m\Phi^{-1}(\gamma_1)\mu}{\xi - 2m\Phi^{-1}(\gamma_1)} \quad \Rightarrow \quad \mathrm{E}[\lambda^*] = \frac{2m_\mu\Phi^{-1}(\gamma_1)\Phi^{-1}(\gamma_2)}{1 - 2m_\mu\Phi^{-1}(\gamma_1)/\xi}$$

A number of the leading comparative statics from the decision-theoretic analysis are maintained. For $\mu > 0$, the non-Duvergerian case leads to strategic switching away from candidate 3. The incentive to vote strategically is once again increasing in the precision of public information (m_μ) and the perceived relative strength of candidate 2 (μ). Unlike the decision-theoretic case, however, a decrease in ξ results in an increase in the strategic incentive. This stems from the self-reinforcing nature of strategic voting. The initial incentive $\lambda_1(0)$ does not depend on ξ. With lower ξ, however, a smaller incentive is required to encourage a voter to act strategically. This further enhances the incentive. A final point is the continuity between the Duvergerian and non-Duvergerian cases. Indeed:

$$\lambda^* \to \infty \quad \text{as} \quad m_\mu \uparrow \frac{\xi}{2\Phi^{-1}(\gamma_1)}$$

Conditioning on accurate signal realisation, the probability that a strategic vote is observed may be calculated. An individual acts strategically whenever $\eta + \epsilon_i + \lambda^* \geq 0 \geq \eta + \epsilon_i$. With

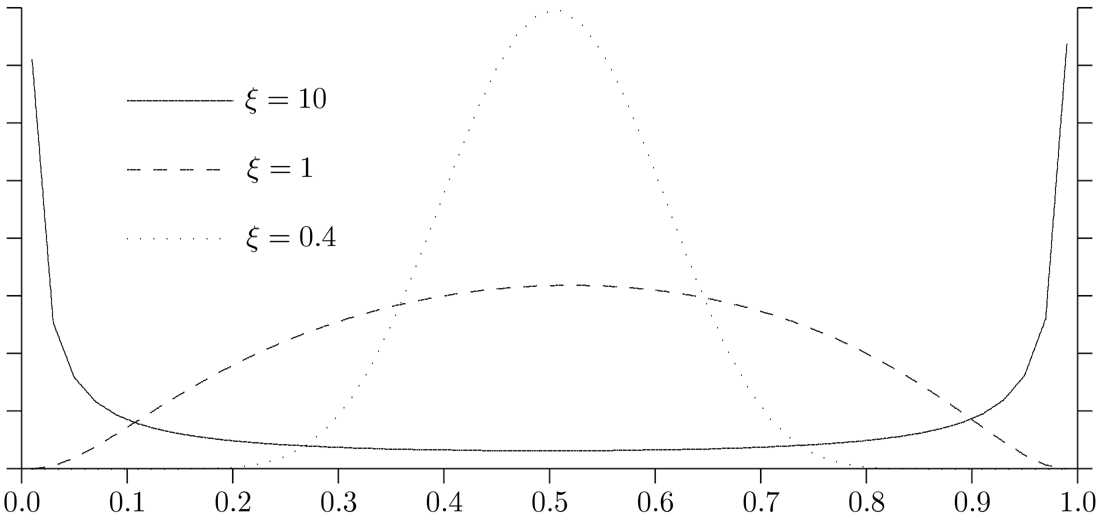


Figure 2.3: Density Over u_{i2} for Various ξ

$\mu = \eta$, this inequality is satisfied with probability:

$$\Pr[\text{Vote } 2 \cup u_{i3} \geq u_{i2} \mid \min\{u_{i2}, u_{i3}\} > u_{i1}] = \Phi\left(\Phi^{-1}(\gamma_2)\left(1 + \frac{2m_\mu \Phi^{-1}(\gamma_1)}{\xi - 2m_\mu \Phi^{-1}(\gamma_1)}\right)\right) - \gamma_2$$

2.5 Illustration

The results of Section 2.4 make it clear that game-theoretic considerations lead to greater strategic voting. How significant is this? This section provides an illustration.

Return to the example of Part 1. The expected vote shares of the three candidates are:

	Candidate	**Share**
1	Conservative	0.38
2	Labour	0.32
3	Liberal Democrat	0.30

Using this specification, $\gamma_1 = 0.38/0.62$ and $\gamma_2 = 0.32/0.62$. Assume for simplicity that the realised signal of the common component equals its true value, so that $\mu = \eta = \xi\Phi^{-1}(\gamma_2)$. There are two parameters remaining. These are the heterogeneity of voter preferences (ξ) and the precision of the public signal (m_μ).

Varying ξ yields a surprisingly rich collection of payoff distributions. To illustrate this, first normalise $u_{i1} = 0$ and $u_{i2} + u_{i3} = 1$. It follows that $0 \leq u_{i2} \leq 1$. This normalisation allows a convenient intepretation of the payoff u_{i2}. Suppose that voter i is indifferent between candidates

1 and 2. This happens whenever $q_2(u_{i2} - u_{i1}) = q_3(u_{i3} - u_{i1})$. Applying the normalisation:

$$\frac{q_3}{q_2 + q_3} = u_{i2}$$

This is the conditional pivotal probability of a LibDem-Con tie that is just sufficient to ensure a vote for the LibDem candidate. It follows that this is a measure of the relative strength of a voter's preference for the Labour candidate.

The distribution over u_2 for various values of ξ is illustrated in Figure 2.3. High heterogeneity corresponds to $\xi = 10$. The distribution over candidate allegiance is clearly bimodal: Voters are either strong Labour or strong Liberal Democrat supporters, and hence require stronger strategic incentives in order to move away from straightforward voting. Contrast this with the case of $\xi = 0.4$. The distribution is unimodal, with most weight towards the centre. This corresponds to near indifference between the Labour and Liberal Democrat candidates.

Figure 2.4 illustrates the probability of observing a strategic vote for the example, using the three values $\xi \in \{0.4, 1, 10\}$. In the public information model, all switching is uni-directional, with voters switching from the Liberal Democrat candidate ($i = 3$) to the Labour candidate ($i = 2$). This means that strategic voting is bounded above by 0.3. The figure considers information precision in the range $0 \leq m_\mu \leq 25$. Strategic voting is clearly increasing in this information precision, and decreasing in heterogeneity ξ. Notice, however, the dramatic effect of moving from a decision-theoretic to a game-theoretic perspective. For instance, in the case of $\xi = 1$, an information precision of only $m_\mu = 1.74$ yields a fully Duvergerian outcome: This is the critical value satisfying $m_\mu = \xi/2\Phi^{-1}(\gamma_1)$. This example suggests that, with only modest information precision, self-reinforcing strategic voting leads to strict bipartism. Such a conclusion, however, is premature. The full analysis of Part 3, including an account of private information sources, is required.

2.6 Conclusion

It is clear that the public information case provides an interesting benchmark. In a non-Duvergerian election, there is uni-directional tactical voting away from the commonly expected trailing candidate. Strategic voting incentives are increasing in the precision of public information, the lead of candidate 1 and the perceived distance between candidates 2 and 3, and decreasing in the idiosyncrasy of voters. With increasing public information, the equilibrium tends to a Duvergerian outcome, achieving this at a *finite* precision level m_μ.

It would be appear from this analysis that the strict version of Duverger's Law — complete local bipartism — might still emerge from a more general model. Indeed, the illustrations of Section 2.5 suggest that only a modest amount of information is required to destablise the non-Duvergerian equilibrium.

Note, however, that the analysis of this part replaces only one assumption of previous models

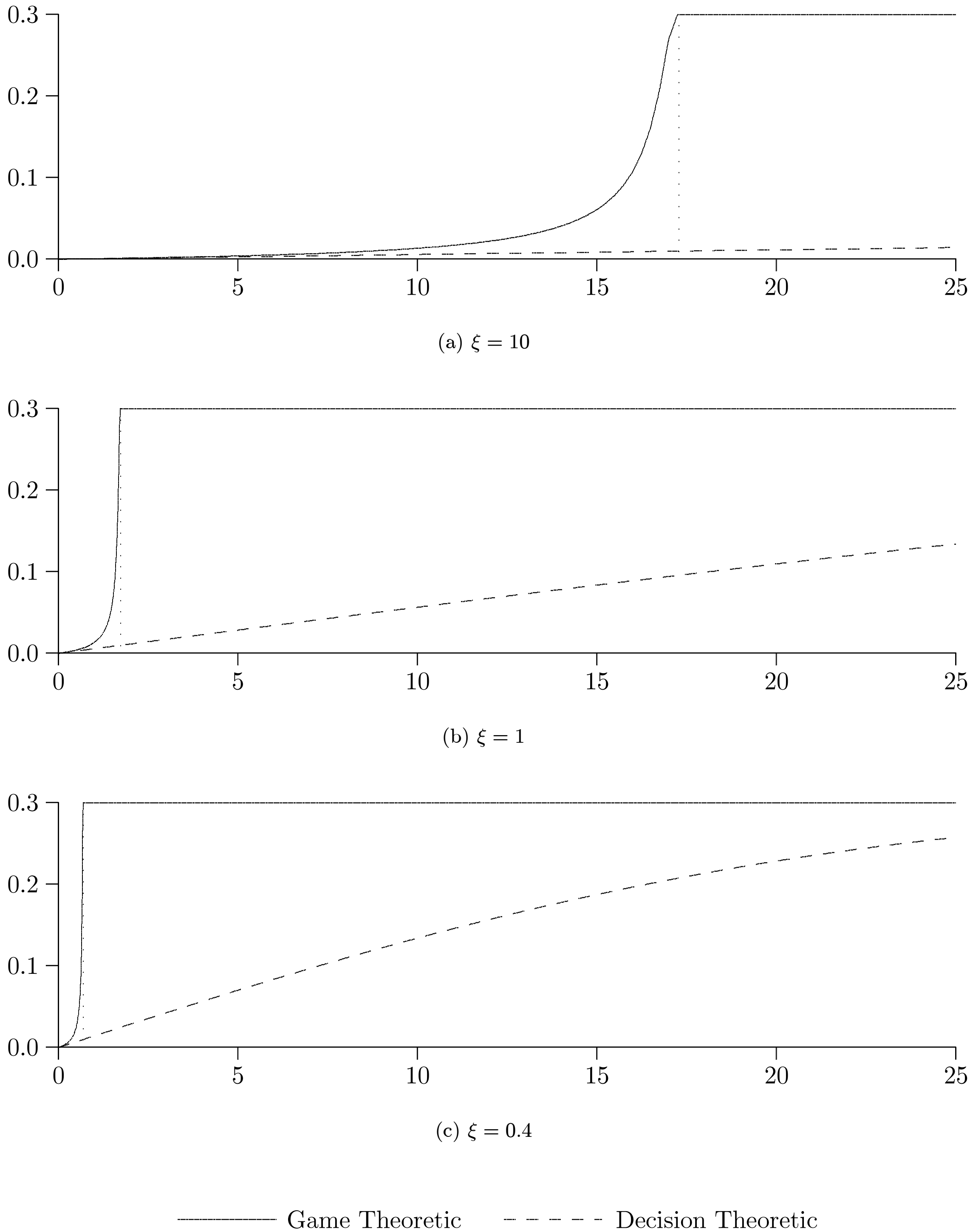


Figure 2.4: Probability of a Strategic Vote from LibDem to Lab for Various m_μ

— the constituency-wide support of candidates 2 and 3 is no longer certain, but all voters hold *common* beliefs. This issue is critical, and is addressed by the analysis of the private signals case, to which the argument now turns.

Part 3

Strategic Voting with Private Signals

"In strategy it is important to see distant things as if they were close and to take a distanced view of close things."

Miyamoto Musashi

3.1 Introduction

Strategic voting does not lead to strict local bipartism. In a decision-theoretic context, this message is clear. Indeed, the analysis of Part 1 demonstrates convincingly that the independence of voter preferences is the key assumption behind the strict version of Duverger's Law. Of course, the game-theoretic analysis of Part 2 appears to resurrect the bipartite outcome. Does it? This part argues not. With public signals as the sole information input to the decision making process, the constituency-wide strength of candidates is uncertain. Voter beliefs, however, remain *common*. It follows that the expected leader is commonly known. This unpalatable feature means that constituency-level uncertainty — the unknown common effect — is insufficient to avoid a Duvergerian outcome. A model is desired in which voters differ in their opinion of which candidate is the likely leader. Indeed, this is a feature that has been recognised in the empirical literature. In their critique of the Niemi *et al* (1992) measure of strategic voting, Heath and Evans (1994) make the following observation:

> "[The Niemi *et al* measure] does not allow for the possibility that some people may intend to avoid wasting their vote, but may be mistaken in their perceptions of the likely chances of the various parties winning the constituency."

This feature is included here, where a rather different equilibrium conclusion is reached when *private* information sources are taken into account.

3.1.1 Self-Attenuating Strategic Voting

Consider once again the standard intuition behind strategic voting. Some rational voters will find it optimal to switch from their preferred candidate to one of the leading pair. This reduces the support for the trailing candidate, and enhances the incentive for others to act strategically. Continuing iteratively, this might (and in Part 2, often does) lead to an equilibrium in which all support for the trailing candidate vanishes. This informal discussion suggests that strategic voting is *self-reinforcing*. But is it? In fact, this part argues that, contrary to immediate intuition, strategic voting may well be *self-attenuating*.

To explore this issue, voters are equipped with two different sources of information. They continue to commonly observe a *public* signal of the common component of preferences — a public signal of constituency-wide support for the candidates. As in Part 2, this may be likened to the observation of an opinion poll. Crucially, each voter also observes a *private* signal, based on the observation of other individuals in the society. Conditional on the common component, this is independent across the electorate.

Part 2 considers only the pure public information case. In that case, strategic voting is indeed self-reinforcing, exhibiting strong positive feedback. If the public signal is sufficiently strong, then the fully Duvergerian equilibria are uniquely stable. The illustrative examples of Section 2.5 suggests that the required information precision is rather small.

With pure private information, this is no longer the case. Strategic voting by others *reduces* the incentive for a voter to act strategically. Why? Essentially, voters fear the prospect of strategic voting in the *opposite* direction to that indicated by their privately-observed signal. In the public information case, this issue is simply not present — all voters observe the same signal, and strategic voting is uni-directional. Intuition is built by considering a voter evaluating the likelihood of either of two challenging candidates tying with a leading candidate. If voters act straightforwardly, then a tie for the lead requires one of the challengers to be significantly stronger than the other. A strategic voter then considers the relative likelihood that each of these challengers has sufficient strength throughout the constituency to force a tie. Suppose instead that other voters act strategically, responding strongly to their private signals of candidate support. Only a small asymmetry in strength between the challengers is necessary to force a tie. A small advantage by one is magnified by strategic voting throughout the electorate. Reflecting on this, a particular voter will weight the relative likelihood of each challenger having a small advantage over the other. This likelihood ratio is less extreme than in the truthful voting case, and hence there is a lower incentive for a voter to act strategically. It follows that strategic voting exhibits *negative* feedback. Indeed, as this informal discussion reveals, the equilibrium incentive for strategic voting in a game-theoretic environment is in fact *less* than that in a decision-theoretic environment. This rules out the possibility of stable Duvergerian equilibria.

The combination of public and private information is the remaining case. The analysis shows that if private information is sufficiently precise relative to public information, then the uniquely stable equilibrium is non-Duvergerian. Which information source is likely to dominate? Certainly, during a national campaign for a legislative election, a large number of public information sources are available. Crucially, however, voting takes place at the *local* level. National opinion polls, for instance, do not reflect local effects, which remain unknown to local constituents. At a constituency level, therefore, one might expect public information to be relatively noisy. Private information dominates, and a non-Duvergerian conclusion emerges. By contrast, if the vote takes place at a national level — for instance, a national referendum or national leadership election — the presence of strong public information sources is likely to dominate. The theory predicts multi-party support at the district-level, and bipartism at the national level.

3.1.2 Comparative Statics

The comparative statics mimic those of Parts 1 and 2. The key difference is the role of information. Private information has a far smaller effect than public information. Indeed, beginning with a Duvergerian election, a non-Duvergerian outcome obtains when sufficient private information is added. It follows that strategic voting may sometimes be *decreasing* in the information available. Perhaps most importantly, in the pure private case, strategic voting is bounded above rather than below by the decision-theoretic level.

3.1.3 Guide

The remainder of this part completes the argument the new theory. Section 3.2 extends the model of Section 2.2 to include privately observed signals. Voter behaviour and the negative feedback inherent in strategic voting is investigated in Section 3.3. Section 3.4 calculates the equilibria and makes a selection. The results are illustrated in Section 3.5, prior to the concluding remarks of Section 3.6. Omitted results are relegated to Section 3.7.

3.2 Model

3.2.1 The Election

The specification of the election is identical to that of Parts 1 and 2. The reader is referred to Section 1.2 on page 5 for full details. There are three candidates $j \in \{1, 2, 3\}$ in a single-seat district selection, operating under the plurality rule. Realised vote totals will be denoted $\{x_j\}$, where $\sum x_j = (1+\gamma_1)n+1$. The vote total for candidate 1 is fixed at $x_1 = \gamma_1 n$, and remaining voters prefer both of candidates 2 and 3 to candidate 1. Assumption 1 is upheld ($\gamma_1 > 1/2$), ensuring that any tie for the lead involves candidate 1. All references to a "voter" and the "electorate" will refer to the $n+1$ remaining voters.

3.2.2 Preference Structure

The preference structure is identical to that adopted in Part 2. Voter i has von Neumann-Morgenstern utility u_{ij} for party j, satisfying:

$$\log \frac{u_{i2} - u_{i1}}{u_{i3} - u_{i1}} = \eta + \epsilon_i \tag{3.1}$$

Here η and ϵ_i are the *common* and *idiosyncratic* components of utility. Maintaining Assumption 3, the idiosyncratic component is distributed independently across voters, with distribution:

$$\epsilon_i \sim N(0, \xi^2)$$

3.2.3 Public and Private Signals

The public signal of Part 2 is augmented by a privately observed signal. Voters begin with a diffuse prior over η — they have no knowledge of average candidate support. Information on candidate support may then be gleaned from a variety of sources. The first is public information, such as the observation of a public opinion poll or similar survey. Assumption 4 is retained, so that voters commonly observe $\mu \sim N(\eta, \sigma^2)$. Following observation of this signal, and prior to the receipt of any private information, voters update to a common public posterior $\eta \sim N(\mu, \sigma^2)$.

Viewed as an opinion poll, Assumption 4 provides a natural framework. In particular, the widespread publication of opinion polls is common during an election. For large m_μ, this leads to a high precision of public information. Notice, however, that opinion polls typically occur at a *national* level. They reflect nationwide support for political parties, but rarely give accurate *constituency* level information. This feature was present in the 1997 British General Election. Evans, Curtice and Norris (1998) note that 47 nationwide opinion polls were conducted during the election campaign. By contrast, only 29 polls were conducted in 26 different constituencies at a constituency level, out of a total of 659 constituencies. To the see the effect of this, decompose the common effect η into national and district components, so that $\eta = \overline{\eta} + (\eta - \overline{\eta})$ where $\overline{\eta}$ is the national component. Even with perfect knowledge of $\overline{\eta}$, uncertainty over the district component remains, leading to a public posterior belief of $\eta \sim N(\mu, \text{var}[\eta - \overline{\eta}])$. At the constituency level, the absence of opinion poll information for the district lowers the precision of public information.

Even in the presence of a constituency level opinion poll, it is perhaps less likely for the poll results to be publicly observed by all members of a constituency. Publicity over such results is often provided by political parties themselves, and these are often manipulated to suit the interests of the candidates.[1] It is thus reasonable to expect the level of information at a

[1] The 1997 British General Election provides an example. In the Oxford West and Abingdon constituency,

constituency level to be rather low. Voters do, however, have other sources of information available. In particular, a signal of constituency-wide candidate support may be obtained from the people with whom an individual interacts. The important characteristic of such information is that it leads to *private* signals.

Assumption 5. *Each voter i observes a* private signal $\delta_i \sim N(\eta, \kappa^2)$. *Conditional on η, this is independent across voters but may be correlated with the idiosyncratic component ϵ_i.*

Once again, a microfoundation is available. The private signal δ_i corresponds to the observation of m_δ randomly chosen individuals, with $\kappa^2 = \xi^2/m_\delta$. In particular, a voter's own payoffs are a signal of η. Hence with $m_\delta = 1$, it follows that $\kappa^2 = \xi^2$. More generally, with this microfoundation, the variance of private signals is bounded above, with $\kappa^2 \leq \xi^2$. The inclusion of a voter's own preferences in the signal results in correlation between the signal and idiosyncratic utility component ϵ_i. For instance, in a sample of $m_\delta > 1$:

$$\delta_i = \eta + \frac{1}{m_\delta}\left\{\epsilon_i + \sum_{k \neq i} \epsilon_k\right\} \quad \Rightarrow \quad \mathrm{cov}(\delta_i, \epsilon_i) = \frac{\xi^2}{m_\delta} = \kappa^2$$

This feature is incorporated into the analysis, and extends easily to further correlation between the preferences of voter i and sampled individuals. Defining the correlation coefficient between δ_i and ϵ_i as ρ, the microfoundation presenteded here yields:

$$\rho \geq \frac{\kappa}{\xi} > 0$$

Following the observation of δ_i, a voter updates to obtain a private posterior belief.

Lemma 6. *The posterior belief of voter i over η satisfies:*

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

Proof: Apply the standard Bayesian updating procedure — see DeGroot (1970). □

The specification of private signals implicitly assumes that sampled individuals reveal their preferences truthfully. Furthermore, since the unconditional probability of influencing the election outcome is vanishingly small, it seems unlikely that individuals would find a costly information acquisition exercise to be worthwhile. The argument presented here accepts this latter critique. If a voter finds it too costly to conduct a private opinion poll, then the strategic manipulation of voting intentions by sampled individuals is no longer of relevance. The question of the private information source remains. It is envisioned that his information is accumulated over an

the "state of play" of the various candidates was publicised by the political parties themselves. The Labour and Liberal Democratic candidates both argued that they were best-placed to take the various seats. Indeed, the author received a variety of conflicting leaflets during the campaign, and eventually decided to ignore their content.

extended period of time prior to an election, in the course of daily activity. It seems unlikely that a sampled individual would find response manipulation worthwhile over such a time frame.

3.3 Voter Behaviour

3.3.1 Optimal Voting

Once again, optimal voting is determined by pivotal considerations. With private information, the perceived pivotal log-likelihood ratio may vary throughout the electorate. Recall Definition 2, and introduce the subscript i so that:

$$\text{Vote } 2 \;\Leftrightarrow\; \log \frac{u_{i2} - u_{i1}}{u_{i3} - u_{i1}} + \lambda_i \geq 0$$

The focus will be on symmetric strategy profiles, and hence symmetric belief rules. It follows that the perceived pivotal log-likelihood ratio depends only on a received private signal δ_i, and not on the voter's identity i. Thus $\lambda_i = \lambda(\delta_i)$. It follows that a voter supports candidate 2 whenever $\eta + \epsilon_i + \lambda(\delta_i) \geq 0$.

3.3.2 Affine Belief Rules under Best Response

Begin by restricting to optimal voting given beliefs. A strategy profile is then entirely determined by the belief rule $\lambda(\delta_i)$. Are there any particular belief rules that are likely to arise in equilibrium? The analysis will reveal the importance of *affine* belief rules.

Definition 5. *An* affine belief rule *satisfies* $\lambda_i = a + b\delta_i$*.*

The parameters a and b may depend on the realisation of the public signal μ as well as the parameters of the model. Notice that truthful voting corresponds to an affine belief rule with $a = b = 0$. Furthermore, Duvergerian equilibria will always be present, obtained by setting $b = 0$ and extending the real line to set $a = \pm\infty$. Focus on affine belief rules is justified by the following lemmata.

Lemma 7. *For an unboundedly large electorate ($n \to \infty$) any smoothly increasing belief rule $\lambda(\delta_i)$ maps to an affine belief rule under best response: When all adopt $\lambda(\delta_i)$, a best response is to adopt an affine belief rule $\hat{\lambda}(\delta_i) = \hat{a} + \hat{b}\delta_i$ for some $\hat{a}$ and $\hat{b}$.*

Proof. See Section 3.7. □

Lemma 8. *The class of affine belief rules is closed under best response. If all adopt a belief* *rule* $\lambda(\delta_i) = a + b\delta_i$*, then a best response is to adopt an affine belief rule* $\hat{\lambda}(\delta_i) = \hat{a} + \hat{b}\delta_i$ *where:*

$$\begin{aligned} \hat{a} = \hat{a}(a,b) &= \frac{\hat{b}[a(\kappa^2+\sigma^2)+(1+b)\kappa^2\mu]}{\sigma^2(1+b)} \\ \hat{b} = \hat{b}(b) &= \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)}{\kappa^2(1+b)} \\ \tilde{\kappa}^2 &= var[\epsilon_i + b(\delta_i - \eta) \,|\, \eta] \end{aligned}$$

Proof. The first part of this lemma is a simple corollary of Lemma 7. The generation of the explict formulae above mimics the proof of Proposition 4, and is described in Section 3.7. □

Following Lemmata 7 and 8, attention will now turn to the properties of affine belief rules. Imposing best response, any rules will remain in this class. First consider the private component b, which describes the response of a voter to the private signal.

Lemma 9. *The mapping* $\hat{b}(b)$ *is strictly decreasing in* b*, with:*

$$\hat{b}(0) = \frac{2\xi\Phi^{-1}(\gamma_1)}{\kappa^2} \quad \textit{and} \quad \lim_{b\to\infty} \hat{b}(b) = \frac{2\Phi^{-1}(\gamma_1)}{\kappa}$$

Proof. See Section 3.7 □

A decision-theoretic conclusion is immediate. Straightforward voting is equivalent to optimal voting with $a = b = 0$. A decision-theoretic voter calculates the optimal voting rule, given straightforward voting by all others. This yields:

$$a_{DT} = \frac{2\xi\Phi^{-1}(\gamma_1)\mu}{\sigma^2} \quad \text{and} \quad b_{DT} = \frac{2\xi\Phi^{-1}(\gamma_1)}{\kappa^2}$$

From this, the pivotal log-likelihood ratio becomes:

$$\lambda_{DT}(\delta_i) = 2\xi\Phi^{-1}(\gamma_1)\left(\frac{\mu}{\sigma^2} + \frac{\delta_i}{\kappa^2}\right)$$

This mirrors the decision-theoretic analysis of Part 2. Attention now turns to the presence of negative feedback with private information.

3.3.3 Self-Attenuating Strategic Voting

Using Lemma 8, notice that $\hat{b}(b)$ is *decreasing* in b. An *increase* in the response to signals by the electorate at large *reduces* the weight placed on the signal by the typical voter: An increase in the tendency to vote strategically by others reduces the incentive for a voter to act strategically. Strategic voting with pure private information is *not* self-reinforcing, in stark contrast to the public information case.

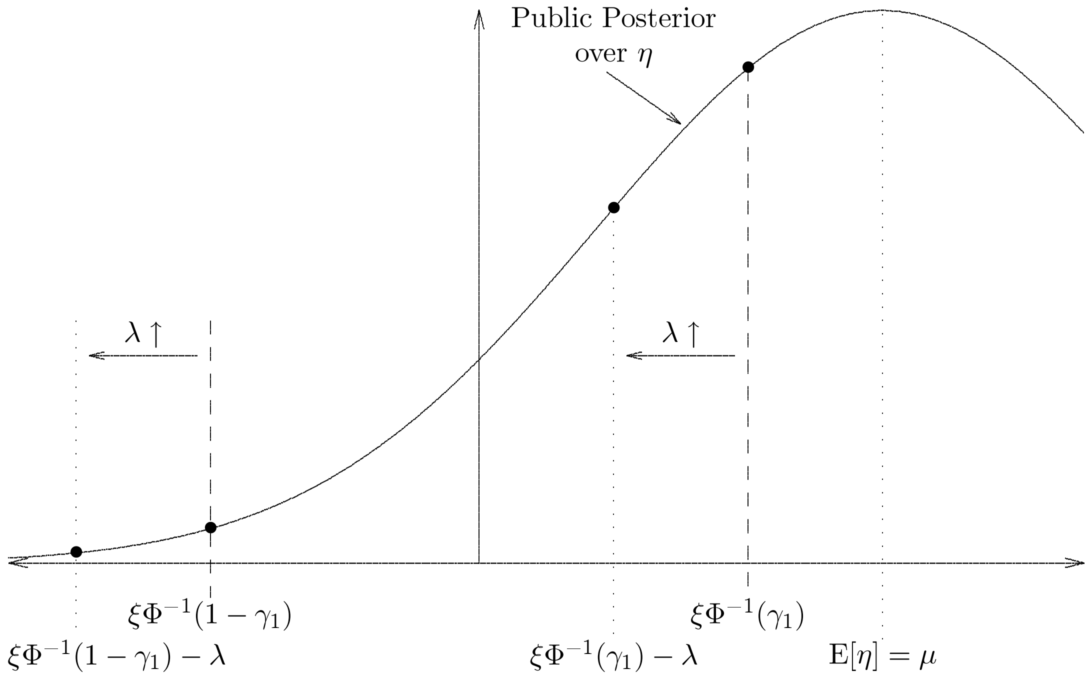


Figure 3.1: Self-Reinforcing Strategic Voting with Public Signals

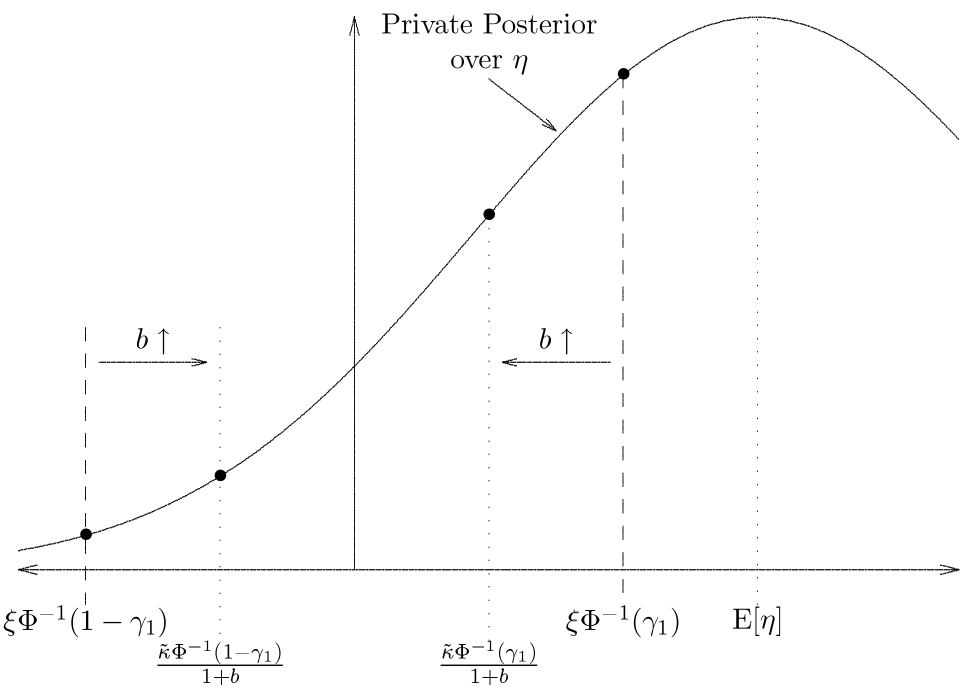


Figure 3.2: Self-Attenuating Strategic Voting with Private Signals

The intuition is aided by Figures 3.1 and 3.2. Consider first the public case. Following observation of the public signal μ, a typical voter holds a common public posterior belief over the common utility component η. This is illustrated in Figure 3.1. A randomly selected individual votes for candidate 2 whenever $\eta+\lambda+\epsilon_i \geq 0$. For a tie with candidate 1, this requires a critical value of $\eta = \xi\Phi^{-1}(\gamma_1) - \lambda$. Similarly, a tie between candidates 1 and 3 requires a critical value of $\eta = \xi\Phi^{-1}(1-\gamma_1) - \lambda$. Increasing λ reduces both of these values and, for $\mu > 0$, increases the likelihood ratio of 1-2 tie versus a 1-3 tie. It follows that strategic voting is self-reinforcing.

By contrast, the private case is illustrated in Figure 3.2. First consider the case of $a = b = 0$. A tie between candidates 1 and 2 requires a critical value of $\eta = \xi\Phi^{-1}(\gamma_1)$. For $b > 0$, a randomly selected individual votes for candidate 2 whenever $\eta + b\delta_i + \epsilon_i \geq 0$, requiring a critical value of $\eta = \tilde{\kappa}\Phi^{-1}(\gamma_1)/(1+b)$. This is decreasing in b. Moving to a 1-3 tie, the critical value of $\eta = \xi\Phi^{-1}(1-\gamma_1)$ increases to $\eta = \tilde{\kappa}\Phi^{-1}(1-\gamma_1)/(1+b)$. This follows, since $\Phi^{-1}(1-\gamma_1) < 0$. The critical values are then closer, so that the likelihood ratio moves closer to 1.

This may at first seem counter-intuitive. When b is high, voters respond strongly to their signal. In particular, this increases the likelihood of a tactical vote. Importantly, it increases the probability of a strategic vote in both directions. It may seem that a voter receiving a signal $\delta_i > 0$ may view the prospect of a 1-3 tie as extremely unlikely. For high δ_i, however, a win by candidate 2 seems almost certain. To affect the outcome, however, a voter *conditions* on the event of a tie. For high δ_i, a tie can *only* occur if the voter's signal has overstated the support of candidate 2. It is then reasonable for the voter to consider true values of η satisfying $\eta < 0$.

Summarising, voters become worried that there may be tactical switching in the opposite direction to that indicated by their private signal.

3.4 Equilibrium

3.4.1 Equilibrium Voting

Once again, there are always Duvergerian equilibria present. Given that all $n+1$ voters support candidate $j \in \{2, 3\}$, the pivotal log-likelihood ratio is undefined, and there is no incentive for voters to switch. A non-Duvergerian equilibrium may be available. Based on the results of the previous section, it is sufficient to examine the class of affine belief rules. A non-Duvergerian equilibrium then corresponds to a finite pair $\{a^*, b^*\}$ such that $b^* = \hat{b}(b^*)$ and $a^* = a(a^*, b^*)$.

The properties of the the mapping $\hat{b}(b)$ immediately yield a unique fixed point.

Lemma 10. *The mapping $\hat{b}(b)$ has a unique fixed point b^* . For $\rho \geq \kappa/\xi$, this satisfies:*

$$\frac{2\Phi^{-1}(\gamma_1)}{\kappa} \leq b^* \leq \frac{\Phi^{-1}(\gamma_1)}{\kappa}\left\{1 + \sqrt{\frac{\Phi^{-1}(\gamma_1) + 2\xi}{\Phi^{-1}(\gamma_1)}}\right\} \tag{3.2}$$

For the microfoundation case, with $\rho = \kappa/\xi$, the bound may be refined to:

$$\frac{2\Phi^{-1}(\gamma_1)}{\kappa} \leq b^* \leq \frac{\Phi^{-1}(\gamma_1)}{\kappa}\sqrt{2 + 2\sqrt{\frac{(\Phi^{-1}(\gamma_1))^2 + \xi^2}{(\Phi^{-1}(\gamma_1))^2}}}$$

Proof. See Section 3.7. □

Any non-Duvergerian equilibrium must entail $b = b^*$. It remains to consider fixed points of $\hat{a}$. Notice that $\hat{a}$ is affine in a, yielding a unique non-Duvergerian equilibrium.

Proposition 8. *For $\sigma^2 \neq b^*\kappa^2$ There is a unique affine non-Duvergerian equilibrium:*

$$a^* = \frac{b^*(1 + b^*)\kappa^2\mu}{\sigma^2 - b^*\kappa^2}$$

Proof. Straightforward solution to $a^* = \hat{a}(a^*, b^*)$. □

Once again, there are multiple equilibria. Moreover, reprising the analysis of Part 2, the direction of the response to the public signal μ is ambiguous, since the denominator of a^* can be either positive or negative.

3.4.2 Equilibrium Selection

An equilibrium selection problem is posed. In the private signal case, there is a unique non-Duvergerian equilibrium, together with two fully Duvergerian equilibria. Stability is once again employed as an equilibrium selection device. Consider an iterative best response process. As in Section 2.4, it is natural to specify an initial hypothesis of straightforward voting. This is trivially an affine belief rule, as it corresponds to $a_0 = b_0 = 0$. Lemma 8 ensures that the best-response process remains within the class of affine belief rules.

Definition 6. *Define the iterative best response process by $b_t = \hat{b}(b_{t-1})$, $a_t = \hat{a}(a_{t-1}, b_{t-1})$.*

The mapping $\hat{b}(b)$ and associated process $\{b_t\}$ is not contingent on a_t, and hence may be considered in isolation.

Lemma 11. *b^* is globally stable in the iterative best response dynamic: $b_t \to b^*$ as $t \to \infty$.*

Proof. See Section 3.7. □

Whereas the formal proof of Lemma 11 is algebraically tedious, a diagrammatic illustration proves useful. Figure 3.3 plots the best response function $\hat{b}(b)$, illustrating the convergence to the fixed point. Notice the cyclic behaviour — this is a consequence of the negative feedback inherent in strategic voting with private signals. Begin with $b_0 = 0$. There is a strong response

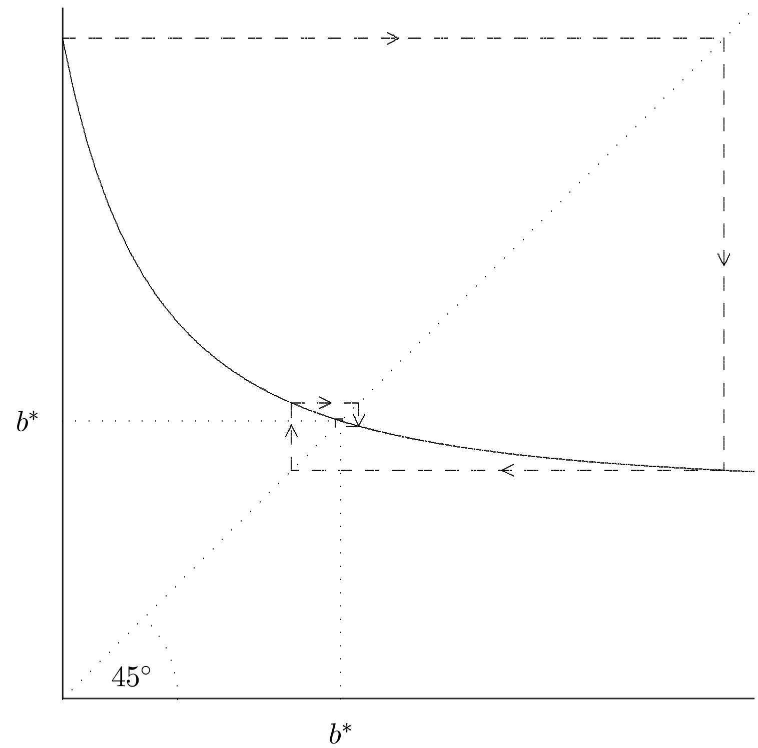


Figure 3.3: $\hat{b}(b)$ and Convergence to Equilibrium

to the private signal — this is the decision theoretic case, yielding $b_1 = b_{DT}$. Taking the next step, the focal voter recognises the strategic behaviour of others. This attenuates the response to the private signal, with a consequent reduction in b. Of course, this behaviour leaves open the possibility of a limit cycle in the iterative best response process. Lemma 11 ensures that the cycling dampens down, eventually converging to the unique fixed point b^*.

To select an equilibrium, now turn to the mapping $\hat{a}(a, b^*)$.

Proposition 9. *If $\sigma^2 > b^*\kappa^2$, then a^* is uniquely stable. A sufficient condition is for:*

$$\sigma^2 > \kappa\Phi^{-1}(\gamma_1)\left\{1 + \sqrt{\frac{\Phi^{-1}(\gamma_1) + 2\xi}{\Phi^{-1}(\gamma_1)}}\right\}$$

If this holds, then the non-Duvergerian equilibrium is uniquely stable, and attained as the limit of the iterative best response process from any finite starting point.

Proof. This follows from the affine nature of the mapping $\hat{a}$. The sufficient condition is obtained by employing the upper bound on b^* from Lemma 10. □

Corollary 1. *With pure private information ($\sigma_2 \to \infty$), the non-Duvergerian equilibrium is uniquely stable, with $a^* = 0$.*

It follows that the non-Duvergerian equilibrium is selected whenever the public information source is sufficiently imprecise. This requirement is also present in the public information model of Part 2. Notice, however, that the precision is judged *relative* to the precision of the private signal.[2] Hence if private signals are relatively more important than public signals, a non-Duvergerian outcome emerges. As Corollary 1 confirms, with only private information, the equilibrium is *always* non-Duvergerian. Which situation is likely to obtain? In a national referendum or similarly nationally conducted election, there are typically many public information sources. Moreover, region-specific effects are unimportant. It follows that public sources are likely to be more important than private. At a district level, however, commonly observed public signals are likely to be fewer. It follows that private signals are the primary source of information. A non-Duvergerian conclusion emerges.

3.4.3 Comparative Statics

The comparative statics for the private information model mimic those of both the decision-theoretic (Part 1) and public information (Part 2) models. In particular, the incentive to vote strategically is increasing in the strength of candidate 1 (γ_1) and the relative strength of candidate 2 over candidate 3 (γ_2). To see this, specialise to the case of pure private information, so that $a^* = 0$ and $\lambda(\delta_i) = b^*\delta_i$. Taking expectations, $\mathrm{E}[\lambda_i] = b^*\xi\Phi^{-1}(\gamma_2)$. Examining the best response function $\hat{b}(b)$:

$$\hat{b} = \frac{2\Phi^{-1}(\gamma_1)\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b}}{\kappa^2(1+b)}$$

The best response (and hence the fixed point b^*) is increasing in γ_1.

There key differences between this and the previous cases. The first is that strategic voting is *bi-directional*; depending on the outcome of the privately observed signal δ_i, a voter may switch in either direction. Indeed, given $\eta > 0$, an "incorrect" signal occurs with probability:

$$\Pr[\delta_i \leq 0] = \Pr\left[\frac{\delta_i - \eta}{\kappa} \leq -\frac{\eta}{\kappa}\right] = 1 - \Phi\left(\frac{\eta}{\kappa}\right) = 1 - \Phi\left(\frac{\xi\Phi^{-1}(\gamma_2)}{\kappa}\right)$$

The second key difference is the effect of information precision on the strategic incentive. To investigate this, specialise to the microfoundation case of $\rho = \kappa/\xi$, and employ the upper bound on b^* from Lemma 10:

$$\mathrm{E}[\lambda_i] = b^*\xi\Phi^{-1}(\gamma_2) \leq \frac{\xi\Phi^{-1}(\gamma_2)\Phi^{-1}(\gamma_1)}{\kappa}\sqrt{2 + 2\sqrt{\frac{\Phi^{-1}(\gamma_1))^2 + \xi^2}{(\Phi^{-1}(\gamma_1))^2}}}$$

[2] In fact, the condition presented here is of the same form as that in Morris and Shin (1999). They consider the coordination problem of bank runs, and find that the ratio of the variance of a public signal and standard deviation of a private signal determines the uniqueness of equilibrium.

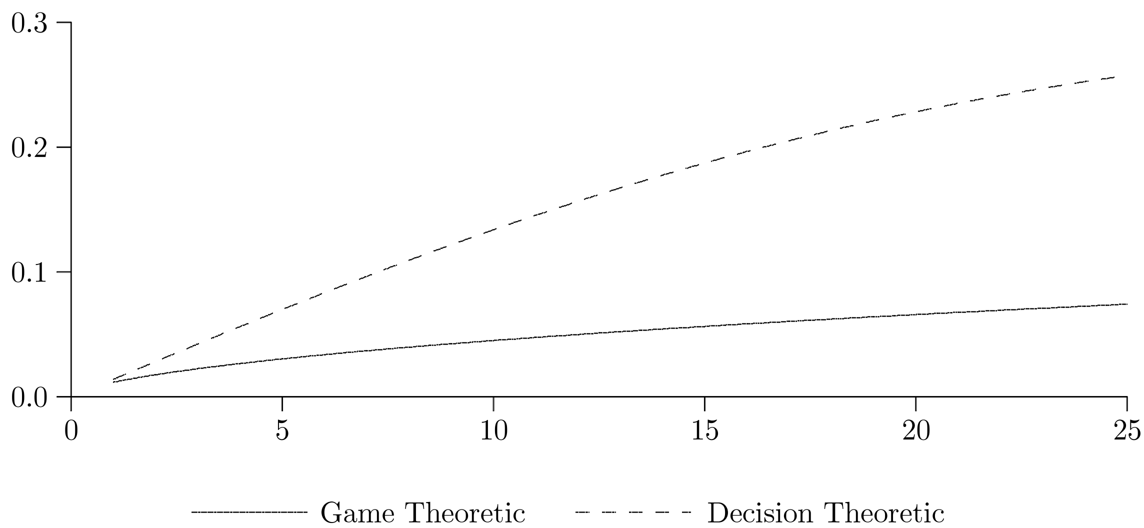


Figure 3.4: Probability of a Switch from 3 to 2 (Private Information, $\xi = 0.4$)

Next, recall that $\xi/\kappa = m_\delta$. Furthermore, allow $\xi \downarrow 0$ to obtain:

$$\mathrm{E}[\lambda_i] \approx 2\Phi^{-1}(\gamma_2)\Phi^{-1}(\gamma_1)\sqrt{m_\delta}$$

There are two pertinent observations. First, as $\xi \downarrow 0$, the expected incentive approaches the lower bound given in Lemma 10. Second, notice the effect of m_δ. In particular, compare this with the decision-theoretic analysis of Part 2 — in particular, Equation (2.5, p. 29). In the decision theoretic case, the strategic incentive is increasing in the precision m. In the private case, however, the strategic incentive is increasing in $\sqrt{m}$. It follows that there is *less* strategic voting than in a decision-theoretic world, and furthermore that the impact of increased signal precision is much less.

3.5 Illustration

A last visit to the example of Part 1 proves useful in illustrating the comparative statics described in Section 3.4. Figure 3.4 illustrates the probability of a strategic switch from candidate 3 to candidate 2. The specification is one of pure private information for $\xi = 0.4$ — it follows that voters are relatively homogeneous, and hence one might expect extensive strategic voting. This is not the case. Notice that strategic voting is significantly below the decision-theoretic benchmark. Negative feedback is the cause. Furthermore, strategic voting from candidate 3 to candidate 2 overstates its effect — there may be individuals who switch in the opposite direction, offsetting the effect. Based on this example, and specifying reasonable precisions for privately observed signal precisions, strategic voting is far from complete — strict local bipartism is rejected.

3.6 Conclusion

The new theory of strategic voting observes that the Cox (1994) and Palfrey (1989) models of strategic voting are driven by the assumption that voter preferences are drawn independently from a commonly-known distribution. This leads to the divergence of the pivotal log-likelihood ratios that are the critical determinants of optimal voting behaviour. Introducing uncertain common effects to voter preferences results in uncertain constituency-wide candidate support, and overturns this result — pivotal log-likelihood ratios remain finite. Moreover, it is *only* uncertainty over common effects that matters. This suggests that the Cox-Palfrey model is perhaps driven by the wrong factors.

The introduction of uncertain common effects, and the careful modelling of voter information sources leads to an elegant model. The key rôle of information is now clear — a fully Duvergerian outcome requires both precise public information and the absence of precise private information. The preferred model is the pure private information model of this part. As Section 3.2 argues, information at a district level is more likely to be private. The model predicts some, but not complete, strategic voting.

Earlier models lack comparative static predictions. This is a necessary consequence of the strict bipartite prediction. Here, the comparative statics are clear. In particular, the model predicts *less* strategic voting in a close election. Strategic voting is also increasing in the homogeneity of the electorate. Evans, Curtice and Norris (1998) hypothesise that increased strategic voting in the 1997 British General Election was likely due to increased indifference between the Labour and Liberal Democrat challengers. This view is consistent with the new theory.

Weaknesses remain. The model considered here specialises to the case of a three-candidate election. Of course, the three-candidate model is of interest in its own right — there are three main challengers in the majority of UK constituencies. More restrictive, however, is the assumption that the votes of candidate 1 are fixed. The introduction of uncertainty over x_1 may further reduce the strategic incentives in the model. To see this, notice that the absolute probability of a pivotal outcome is larger as the strength of candidate 1 vanishes. This will likely reduce the pivotal likelihood ratio of ties between the possible tying pairs. Perhaps more serious is the restriction on voter preference orderings. The structure of preferences rules out multi-direction strategic switching; all n voters are assumed to support either candidate 2 or candidate 3. In the case of $\eta > 0$, this ensures that candidate 2 is in fact a Condorcet winner. This restriction requires lifting.

All of these issues are the subject of ongoing research. In addition, empirical testing of the model is already in progress, and experimental work is planned. Hopefully a greater understanding of strategic behaviour may lead to a better understanding of electoral systems. Alternatively, the reader may wish to adopt the sentiments of Winston Churchill, who claimed that "[t]he best argument against democracy is a five minute conversation with the average voter." Perhaps a conversation with the pivotal voter would have confirmed his views.

3.7 Omitted Results

3.7.1 Properties of $\hat{b}(b)$

This section analyses the properties of the best response function $\hat{b}(b)$ described in Section 3.3.

Proof of Lemma 7: Consider an arbitrary smoothly increasing belief rule $\lambda(\delta_i)$. Define:

$$p = \Pr[\lambda(\delta_i) + \eta + \epsilon_i \geq 0 \,|\, \eta] = H(\eta)$$

This is the probability that a randomly selected voter supports candidate 2, given the common utility component η. Given that $\lambda(\delta_i)$ is smoothly increasing, $H(\eta)$ is strictly increasing in η. Write $h(\eta) = H'(\eta)$. It follows that:

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

This probability (via $\mathrm{E}(\eta)$ and $\mathrm{var}[\eta]$) is conditional on the information available to a focal voter, and uses the fact that posterior beliefs over η are normal. Differentiate to obtain:

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

Next, using Proposition 1 from Part 1, recall that:

$$\lim_{n\to\infty}\frac{q_2}{q_3} = \frac{f(\gamma_1)}{f(1-\gamma_1)}$$

Evaluate this expression and take logs to obtain:

$$\begin{aligned}\log\frac{f(\gamma_1)}{f(1-\gamma_1)} &= \log\frac{h(H^{-1}(1-\gamma_1))}{h(H^{-1}(\gamma_1))} - \frac{(H^{-1}(\gamma_1) - \mathrm{E}[\eta])^2}{2\mathrm{var}[\eta]} + \frac{(H^{-1}(1-\gamma_1) - \mathrm{E}[\eta])^2}{2\mathrm{var}[\eta]} \\ &= \log\frac{h(H^{-1}(1-\gamma_1))}{h(H^{-1}(\gamma_1))} + \frac{H^{-1}(1-\gamma_1)^2 - H^{-1}(\gamma_1)^2}{2\mathrm{var}[\eta]} + \frac{H^{-1}(\gamma_1) - H^{-1}(1-\gamma_1)}{\mathrm{var}[\eta]}\mathrm{E}[\eta] \qquad (3.3)\end{aligned}$$

Notice that this is affine in $\mathrm{E}[\eta]$. Next recall from Lemma 6 of Section 3.2 that:

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

Thus $\mathrm{var}[\eta]$ does not depend on δ_i, and $\mathrm{E}[\eta]$ is affine in δ_i, so that $\hat{\lambda}(\delta_i)$ is affine in δ_i. $\square$

Proof of Lemma 8: Suppose that each voter adopts the initial affine belief rule. This results in a voter for candidate 2 whenever:

$$\eta + a + b\delta_i + \epsilon_i \geq 0 \;\;\Leftrightarrow\;\; a + (1+b)\eta \geq -\epsilon_i - b(\delta_i - \eta)$$

Conditional on η, the right hand side is normally distributed with zero expectation, and variance $\tilde{\kappa}^2 = \mathrm{var}[\epsilon_i + b(\delta_i - \eta)]$. It follows that:

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

Differentiate to obtain:

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

Begin with the first term of Equation (3.3). First employ the symmetry of the normal distribution to note that $\Phi^{-1}(1-\gamma_1) = -\Phi^{-1}(\gamma_1)$, and that $\phi(z) = \phi(-z)$. It follows that:

$$\log\frac{h(H^{-1}(1-\gamma_1))}{h(H^{-1}(\gamma_1))} = \log\frac{\phi(\Phi^{-1}(1-\gamma_1))}{\phi(\Phi^{-1}(\gamma_1))} = 0$$

Next consider the second term of Equation (3.3).

$$H^{-1}(\gamma_1)^2 = \frac{(\tilde{\kappa}\Phi^{-1}(\gamma_1) - a)^2}{(1+b)^2} = \frac{[\tilde{\kappa}\Phi^{-1}(\gamma_1)]^2 + a^2 - 2a\tilde{\kappa}\Phi^{-1}(\gamma_1)}{(1+b)^2}$$

Similarly:

$$H^{-1}(1-\gamma_1)^2 = \frac{[\tilde{\kappa}\Phi^{-1}(1-\gamma_1)]^2 + a^2 - 2a\tilde{\kappa}\Phi^{-1}(1-\gamma_1)}{(1+b)^2} = \frac{[\tilde{\kappa}\Phi^{-1}(\gamma_1)]^2 + a^2 + 2a\tilde{\kappa}\Phi^{-1}(\gamma_1)}{(1+b)^2}$$

It follows that:

$$\frac{H^{-1}(1-\gamma_1)^2 - H^{-1}(\gamma_1)^2}{2\mathrm{var}[\eta]} = \frac{2a\tilde{\kappa}\Phi^{-1}(\gamma_1)}{\mathrm{var}[\eta](1+b)^2}$$

The final term is simply:

$$\frac{H^{-1}(\gamma_1) - H^{-1}(1-\gamma_1)}{\mathrm{var}[\eta]}\mathrm{E}[\eta] = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)(1+b)}{(1+b)^2\mathrm{var}[\eta]}\mathrm{E}[\eta]$$

Assembling, obtain:

$$\log\frac{f(\gamma_1)}{f(1-\gamma_1)} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)(a + (1+b)\mathrm{E}[\eta])}{\mathrm{var}[\eta](1+b)^2}$$

The next step is the evaluation of the expectation and variance of η, conditional on the information of a typical voter. Once again, recalling Lemma 6 of Section 3.2 it follows that:

$$\mathrm{E}[\eta] = \frac{\kappa^2\mu + \sigma^2\delta_i}{\kappa^2 + \sigma^2} \quad \text{and} \quad \mathrm{var}[\eta] = \frac{\kappa^2\sigma^2}{\kappa^2 + \sigma^2}$$

Substitute in to obtain:

$$\hat{\lambda}(\delta_i) = \log \frac{f(\gamma_1)}{f(1-\gamma_1)} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)(a(\kappa^2+\sigma^2)+(1+b)(\kappa^2\mu+\sigma^2\delta_i))}{\kappa^2\sigma^2(1+b)^2}$$

Separate this out to obtain the affine function:

$$\hat{\lambda} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)[a(\kappa^2+\sigma^2)+(1+b)\kappa^2\mu]}{\kappa^2\sigma^2(1+b)^2} + \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)\delta_i}{\kappa^2(1+b)}$$

Taking the intercept:

$$\begin{aligned}
\frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)[a(\kappa^2+\sigma^2)+(1+b)\kappa^2\mu]}{\kappa^2\sigma^2(1+b)^2} &= \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)\sigma^2}{\kappa^2\sigma^2(1+b)} \frac{[a(\kappa^2+\sigma^2)+(1+b)\kappa^2\mu]}{\sigma^2(1+b)} \\
&= \frac{\hat{b}[a(\kappa^2+\sigma^2)+(1+b)\kappa^2\mu]}{\sigma^2(1+b)}
\end{aligned}$$

This yields the desired result. □

Proof of Lemma 9: From Lemma 8 the best response $\hat{b}(b)$ satisfies:

$$\hat{b} = \frac{2\tilde{\kappa}\Phi^{-1}(\gamma_1)}{\kappa^2(1+b)}$$

This requires evaluation of $\tilde{\kappa}$, which satisfies:

$$\tilde{\kappa}^2 = \mathrm{var}[\epsilon_i + b(\delta_i - \eta) \,|\, \eta] = \xi^2 + b^2\kappa^2 + 2\rho\kappa\xi b$$

where ρ is the correlation coefficient between ϵ_i and δ_i, conditional on the common utility component η. Hence:

$$\hat{b} = \frac{2\Phi^{-1}(\gamma_1)}{\kappa^2} \frac{\sqrt{\xi^2+b^2\kappa^2+2\rho\kappa\xi b}}{1+b}$$

Evaluating the derivative:

$$\begin{aligned}
\hat{b}'(b) &= \frac{2\Phi^{-1}(\gamma_1)}{\kappa^2}\left\{\frac{b\kappa^2+\rho\kappa\xi}{(1+b)\sqrt{\xi^2+b^2\kappa^2+2\rho\kappa\xi b}} - \frac{\sqrt{\xi^2+b^2\kappa^2+2\rho\kappa\xi b}}{(1+b)^2}\right\} \\
&= \frac{2\Phi^{-1}(\gamma_1)\sqrt{\xi^2+b^2\kappa^2+2\rho\kappa\xi b}}{\kappa^2(1+b)}\left\{\frac{b\kappa^2+\rho\kappa\xi}{\xi^2+b^2\kappa^2+2\rho\kappa\xi b} - \frac{1}{1+b}\right\} \\
&= \hat{b}(b)\left\{\frac{b\kappa^2+\rho\kappa\xi}{\xi^2+b^2\kappa^2+2\rho\kappa\xi b} - \frac{1}{1+b}\right\}
\end{aligned}$$

This is decreasing for $b \geq 0$ if:

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

Re-arrange this expression to obtain $\xi(\rho\kappa - \xi) \leq b\kappa(\rho\xi - \kappa)$ To check this inequality, first consider the right hand side. First $\rho \geq \kappa/\xi$ by assumption — see Section 3.2. Since $b \geq 0$, it is sufficient to show that the left hand side is weakly negative, which requires $\xi \geq \rho\kappa$. But this holds, since $0 \leq \rho \leq 1$ and $\kappa \leq \xi$. It follows that the function is (weakly) decreasing everywhere. Next, evaluate at the extremes to obtain:

$$\hat{b}(0) = \frac{2\xi\Phi^{-1}(\gamma_1)}{\kappa^2} \quad \text{and} \quad \lim_{b\to\infty} \hat{b}(b) = \frac{2\Phi^{-1}(\gamma_1)}{\kappa}$$

These calculations yield the desired properties of the function. □

Proof of Lemma 10: To obtain an upper bound for the fixed point b^*, write $\hat{b}(b)$ as:

$$\hat{b}(b)\kappa = \frac{2\Phi^{-1}(\gamma_1)\sqrt{\xi^2 + b^2\kappa^2 + 2\rho\xi\kappa b}}{\kappa + b\kappa}$$

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

$$\hat{\beta}(\beta) = \frac{2\Phi^{-1}(\gamma_1)\sqrt{\xi^2 + \beta^2 + 2\rho\xi\beta}}{\kappa + \beta} \leq \frac{2\Phi^{-1}(\gamma_1)(\xi + \beta)}{\beta}$$

An upper bound may now be obtained by solving the equation:

$$\beta^2 - 2\Phi^{-1}(\gamma_1)(\xi + \beta) = 0$$

This has a positive root at:

$$\beta = \Phi^{-1}(\gamma_1)\left\{1 + \sqrt{\frac{\Phi^{-1}(\gamma_1) + 2\xi}{\Phi^{-1}(\gamma_1)}}\right\}$$

It follows that an upper bound for the fixed point is:

$$b^* \leq \frac{\Phi^{-1}(\gamma_1)}{\kappa}\left\{1 + \sqrt{\frac{\Phi^{-1}(\gamma_1) + 2\xi}{\Phi^{-1}(\gamma_1)}}\right\}$$

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

$$\hat{\beta}(\beta) = \frac{2\Phi^{-1}(\gamma_1)\sqrt{\xi^2 + \beta^2 + 2\kappa\beta}}{\kappa + \beta} = 2\Phi^{-1}(\gamma_1)\sqrt{\frac{\xi^2 + \beta^2 + 2\kappa\beta}{\kappa^2 + \beta^2 + 2\kappa\beta}}$$

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

$$\hat{\beta}(\beta) \leq \frac{2\Phi^{-1}(\gamma_1)\sqrt{\xi^2 + \beta^2}}{\beta}$$

To obtain an upper bound, solve the equation:

$$\beta^4 - (2\Phi^{-1}(\gamma_1))^2(\xi^2 + \beta^2) = 0$$

This equation is quadratic in β^2, and may be solved to obtain the positive root:

$$\begin{aligned}\beta^2 &= \frac{(2\Phi^{-1}(\gamma_1))^2 + \sqrt{(2\Phi^{-1}(\gamma_1))^4 + 4(2\Phi^{-1}(\gamma_1))^2\xi^2}}{2} \\ &= \frac{(2\Phi^{-1}(\gamma_1))^2}{2}\left\{1 + \sqrt{1 + \frac{\xi^2}{(\Phi^{-1}(\gamma_1))^2}}\right\}\end{aligned}$$

It follows that an upper bound is:

$$b^* \leq \frac{\Phi^{-1}(\gamma_1)}{\kappa}\sqrt{2 + 2\sqrt{1 + \frac{\xi^2}{(\Phi^{-1}(\gamma_1))^2}}}$$

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

3.7.2 Stability of $\hat{b}(b)$

Proof of Lemma 11: Consider the mapping:

$$B(b) = \hat{b}^{(2)}(b) = \hat{b}(\hat{b}(b))$$

Notice that $\hat{b}$ is also a fixed point of B. Taking the derivative of this function:

$$B'(b) = \hat{b}'(\hat{b}(b))\hat{b}'(b)$$

It follows that this is an increasing function, since $\hat{b}' \leq 0$. Consider a generic fixed point b, satisfying $B(b) = b$. Evaluate the derivative at this fixed point:

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

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

Bibliography

Butler, D., and D. Kavanagh (1997): *The British General Election of 1997*. MacMillan, London.

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

Casella, G., and R. L. Berger (1990): *Statistical Inference*. Duxbury Press, Belmont, CA.

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

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

Curtice, J., and M. Steed (1997): "The Results Analysed," in *The British General Election of 1997*, ed. by D. Butler, and D. Kavanagh. MacMillan, London.

DeGroot, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, 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.

Dworak, J. A., J. P. Johnson, and D. P. Myatt (1999): "Team Play in the War of Attrition," mimeo, Nuffield College, Oxford and MIT Department of Economics.

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.

Farquharson, R. (1969): *Theory of Voting*. Yale University Press, New Haven.

Fey, M. (1997): "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting," *American Political Science Review*, 91(1), 135–147.

Heath, A., and G. Evans (1994): "Tactical Voting: Concepts, Measurements and Findings," *British Journal of Political Science*, 24, 557–561.

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

Johnston, R. J., and C. J. Pattie (1991): "Tactical Voting in Great Britain in 1983 and 1987: An Alternative Approach," *British Journal of Political Science*, 21, 95–108.

Lanoue, D. J., and S. Bowler (1992): "The Sources of Tactical Voting in British Parliamentary Elections, 1983-1987," *Political Behaviour*, 14, 141–157.

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

Morris, S., and H. S. Shin (1999): "Coordination Risk and the Price of Debt," mimeo Nuffield College, Oxford.

Myatt, D. P., and C. Wallace (1997): "Adaptive Dynamics with Idiosyncratic Agents," Nuffield College Economics Discussion Paper W31.

——— (1999): "Sophisticated Play by Idiosyncratic Agents," Nuffield College Economics Discussion Paper 1999-W1.

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

Niemi, R., G. Whitten, and M. Franklin (1992): "Constituency Characteristics, Individual Characteristics and Tactical Voting in the 1987 British General Election," *British Journal of Political Science*, 22, 229–240.

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

Riker, W. H. (1976): "The Number of Political Parties: A Reexamination of Duverger's Law," *Comparative Politics*, 9, 753–766.

——— (1982): "The Two-Party System and Duverger's Law: An Essay on the History of Political Science," *American Political Science Review*, 76(4), 753–766.