

Center for Mathematical for Studies in Economics and Management Science

Northwestern University, Evanston, Illinois 60208

Internet: <http://www.kellogg.nwu.edu/research/math/>

Discussion Paper No. 1214

COMPARISON OF SCORING RULES IN POISSON VOTING GAMES

by

Roger B. Myerson*

July 1998

revised, May 1999

Abstract. Scoring rules are compared by the equilibria that they generate for simple elections with three candidates and voters drawn from large Poisson distributions. A calculus for comparing pivot probabilities in Poisson voting games is applied. For a symmetric Condorcet cycle, nonsymmetric discriminatory equilibria exist under best-rewarding scoring rules like plurality voting. A candidate who is universally disliked may still not be out of contention under worst-punishing scoring rules like negative-plurality voting. In elections where two of three candidates have the same position, symmetric equilibria coincide with majority rule only for scoring rules that are balanced between best-rewarding and worst-punishing. When voters also care about continuous functions of vote shares, equilibria may still depend on pivot probabilities.

*Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208-2009.
E-mail: myerson@nwu.edu

COMPARISON OF SCORING RULES IN POISSON VOTING GAMES

by Roger B. Myerson

1. From social choice theory to analysis of voting games

The great impossibility theorems of social choice theory teach us that no democratic voting rule can guarantee the existence of a unique pure-strategy equilibrium in all social choice situations (Muller and Satterthwaite, 1977). Thus, any democratic voting rule must sometimes generate randomized equilibria or multiple equilibria in which the chosen alternative will not depend uniquely on the voters' preferences. But such impossibility results do not imply that all voting rules generate the same sets of equilibria for a given social choice situation. In fact, the set of equilibrium outcomes in a social choice situation (where voters' preferences over the alternatives are held fixed) may depend substantially on the voting rule. To develop a practical theory of social choice and constitution design, we need to understand this dependence. That is, we need to move from the negative impossibility theorems to a positive research agenda of characterizing and comparing the equilibria that are generated by different voting rules. In this paper, we consider some simple social choice situations involving just three alternatives, and with these examples we probe the ways that sets of equilibria may change when the voting rule is changed.

The example that is considered first and longest in this paper is a version of the Condorcet cycle. Beginning with this example can help us to see more clearly the relationship between the impossibility theorems of social choice theory and the analysis of voting games, because the simple Condorcet cycle offers the easiest way to prove a social-choice impossibility theorem for

anonymous neutral voting rules. In social choice theory, a voting rule is called anonymous if it treats the voters symmetrically, and it is called neutral if it treats the candidates symmetrically. Such anonymity and neutrality may be assumed as basic properties of any democratic voting rule.

In the simplest version of the Condorcet cycle, there are three candidates (or social-choice alternatives) which we may number 1, 2, and 3, and there are three voters whose preference orderings are $1 \succ 2 \succ 3$, $2 \succ 3 \succ 1$, and $3 \succ 1 \succ 2$, so that each candidate could be beaten by a majority that prefers another candidate. This simple Condorcet cycle has obvious symmetries among the voters and candidates. So when we complete the definition of a voting game for this example by specifying an anonymous neutral voting rule, the three candidates must appear symmetrically in the resulting set of Nash equilibria. If the voting game for this example has a pure-strategy equilibrium in which candidate 1 would win the election for sure, then by symmetry it must also have a second equilibrium in which candidate 2 would win the election, and it must also have a third equilibrium in which candidate 3 would win the election.

Thus, under any anonymous neutral voting rule, the simple Condorcet cycle cannot have a unique equilibrium in which one candidate wins without randomization. That is, under any democratic voting rule, the Condorcet cycle will confront us with problems of randomization or multiplicity of equilibria. If a voting rule yields here a unique Nash equilibrium, then it must be a randomized equilibrium in which the winner of the election is unpredictable, depending on random factors that are not publicly known before election day. On the other hand, if we have multiple equilibria, then the winner of the election may be predictable once we know which equilibrium is focal in the voters' perceptions, but this focal equilibrium selection may depend crucially on payoff-irrelevant factors of cultural tradition or heresthetic manipulation.

This argument suggests that it may be useful to distinguish among voting rules on the basis of whether they confront us with randomization or multiplicity of equilibria in examples like the Condorcet cycle. But the small number of voters in this simple Condorcet cycle may make it less attractive as a test case for categorizing voting rules that are intended for elections with large numbers of voters. So in this paper we consider instead a replicated version of this example where the voting population can be much larger, including many voters like each of the three voters in the simple version of the Condorcet cycle. Then, to complement and extend the insights from this Condorcet-cycle example, we go on to consider three other simple social choice situations.

A social choice situation here is defined by specifying a set of candidates (or social-choice alternatives) and a probabilistic description of how many voters will have each possible utility function over this set of candidates. To complete the definition of a voting game, we must additionally specify the voting rule that defines what permissible ballots can be cast by the voters and how the winning candidate will be determined by the voters' ballots.

In this paper, the voting rules that we consider are scoring rules. For three-candidate elections, a scoring rule is characterized by a set $C \subseteq \mathbb{R}^3$ which represents the set of feasible ballots or vote-vectors that a voter is permitted to submit. In this set C , a vote-vector $c=(c_1,c_2,c_3)$ represents a ballot that gives c_1 points to candidate 1, c_2 points to candidate 2, and c_3 points to candidate 3. At the end of the election, the vote-vectors of all the voters are added together to compute a total point score for each candidate, and the winner of the election will be the candidate who gets the highest score. In the event of a tie for the highest score, we assume that the winner will be selected randomly from the set of tied candidates, each with equal probability.

We assume throughout this paper that the set C is a nonempty finite subset of \mathbb{R}^3 and that every component c_i of every vector c in C is a rational number. In plurality voting, for example, the set of feasible vote-vectors is

$$C = \{(1,0,0), (0,1,0), (0,0,1), (0,0,0)\}$$

because each voter must vote for exactly one candidate or abstain. In approval voting, each voter can give 0 or 1 point to each candidate, independently of the number of points that he has given to other candidates, and so the set of feasible approval vote-vectors is

$$C = \{(1,0,0), (0,1,0), (0,0,1), (0,1,1), (1,0,1), (1,1,0), (0,0,0)\}.$$

(We can ignore the option of approving everyone, which is equivalent to abstaining.) In Borda voting, voters must rank-order the candidates and assign 1 point to the top-ranked candidate, 1/2 point to the middle candidate, and 0 points to the bottom candidate, and so the set of feasible Borda vote-vectors is

$$C = \{(1,1/2,0), (1,0,1/2), (1/2,1,0), (0,1,1/2), (1/2,0,1), (0,1/2,1), (0,0,0)\}.$$

Negative-plurality voting (in which each voter votes against one candidate) may be represented by the feasible set

$$C = \{(1,1,0), (1,0,1), (0,1,1), (0,0,0)\}.$$

Here $(1,1,0)$ can be interpreted as a vote against candidate 3.

All these scoring rules are anonymous, because we are assuming that every voter can choose from the same feasible set C . The scoring rules that we consider in this paper are also assumed to be neutral in the sense that, for any one-to-one permutation $\zeta: \{1,2,3\} \rightarrow \{1,2,3\}$ on the set of candidates and for any feasible vote-vector (c_1, c_2, c_3) in C , the vector $(c_{\zeta(1)}, c_{\zeta(2)}, c_{\zeta(3)})$ is also a feasible vote-vector in C . We will also assume (without loss of generality) that the scoring rules

are normalized so that the maximal difference that one voter can make between any two candidates is one point; that is, for any two candidates i and j ,

$$\max_{c \in C} c_i - c_j = 1,$$

and that the vote-vectors that achieve this maximum have $c_i = 1$ and $c_j = 0$.

Before we can analyze the equilibria of examples like the large replicated Condorcet cycle, we need some basic concepts and technical results. Section 2 develops the general models and solution concepts which are used in this paper. In particular, within a set of perfect asymptotic equilibria, we look for discriminatory equilibria in which the voters' perceptions of the candidates are very nonsymmetrical. Section 3 then presents some basic techniques that we will use to evaluate the probabilities of pivot events in large Poisson voting games, applying and extending results from Myerson (1999). (New proofs are in Section 8.) With these results and definitions, we characterize the scoring rules that generate discriminatory equilibria for the Poisson-replicated version of the Condorcet cycle in Section 4. We find that multiple discriminatory equilibria exist for scoring rules that put more emphasis on rewarding the top-ranked or "best" candidate (as in plurality voting), rather than on punishing the bottom-ranked or "worst" candidate (as in negative-plurality voting).

Discriminatory equilibria seem problematic in the Condorcet cycle. Section 5 analyzes the discriminatory equilibria of a second simple example where the existence of a discriminatory equilibrium seems highly desirable. For this example, we find that a discriminatory equilibrium fails to exist for scoring rules that put more emphasis on punishing the worst candidate rather than rewarding the best. So the scoring rules that appear best for our first two examples are those that are intermediate between best-rewarding and worst-punishing, especially those that offer voters a

choice between these two forms of expression.

In Section 6 we switch to looking at symmetric equilibria. Applying methods of Cox (1987, 1990) and Myerson (1993), we consider an example where two of the three candidates are indistinguishable from the perspective of the voters, and we show that the symmetric equilibria are consistent with majority preference only for scoring rules that are exactly balanced between best-rewarding and worst-punishing.

The methodology of Sections 2 through 4 put heavy emphasis on calculating probabilities that one vote could change the discrete winner of the election. A reader might suspect that rational voters would ignore these small pivot probabilities when they also consider the continuous effects that a change in relative vote shares may have on subsequent political decisions. In Section 7, we consider a simple example from Piketty (1995) which shows that continuous vote-share effects do not necessarily overwhelm discrete pivot effects.

2. Pivot probabilities and discriminatory equilibria

In this paper, we consider social choice situations where the number of voters is a random variable with a large mean, and the voters have independent private values over the candidates or alternatives. So we start by specifying the set of candidates or alternatives, which here (until Section 7) is the set $\{1,2,3\}$. The number of voters is assumed to be a Poisson random variable with mean n (as in Myerson, 1999). To study large games, we will consider a sequence of games parameterized by n and look at the limit as $n \rightarrow \infty$. In this paper, we will assume that voters differ only in their independent private values for the various candidates. Thus each type of voter here can be characterized by a vector that lists the voter's utilities for the various candidates. That is,

we may suppose that a voter's type t is a triple

$$t = (t_1, t_2, t_3)$$

where each t_i is the utility payoff that the voter would get if candidate i won the election. We assume throughout this paper that the set of possible types T is a nonempty finite subset of \mathbb{R}^3 .

We also assume that each voter's type is independently drawn from some probability distribution r over this set. For example, in Section 4 we will consider a Poisson-replicated version of the Condorcet cycle where the type set is

$$T = \{(10,5,0), (0,10,5), (5,0,10)\}$$

and the probability distribution r is

$$r(10,5,0) = r(0,10,5) = r(5,0,10) = 1/3,$$

and so any voter is equally likely to have the preference ordering $1 \succ 2 \succ 3$ or $2 \succ 3 \succ 1$ or $3 \succ 1 \succ 2$.

Then we complete the definition of a Poisson voting game by specifying the set $C \subseteq \mathbb{R}^3$ which represents the set of feasible ballots or vote-vectors $c = (c_1, c_2, c_3)$ that a voter is permitted to submit in the election.

Given this set of choices C , beliefs about aggregate voting behavior can be described by a probability distribution $\tau = (\tau(c))_{c \in C}$, where $\tau(c)$ denotes the probability that a randomly sampled voter will cast the vote-vector c . This distribution must be in the set of all probability distributions on C ,

$$\Delta(C) = \{\tau \in \mathbb{R}^C \mid \tau(c) \geq 0 \forall c \in C, \sum_{d \in C} \tau(d) = 1\}.$$

The results of the voter's aggregate behavior will be some vote profile $x = (x(c))_{c \in C}$ where $x(c)$ is a nonnegative integer denoting the number of voters who cast the vote-vector c . We let $Z(C)$ denote the set of all such vote profiles,

$$Z(C) = \{x \in \mathbb{R}^C \mid x(c) \text{ is a nonnegative integer } \forall c \in C\}.$$

If the probability distribution τ_n in $\Delta(C)$ describes beliefs about the behavior of any randomly sampled voter in a game with expected population size n , then the vector $n\tau_n = (n\tau_n(c))_{c \in C}$ is the expected vote profile. If the expected number of voters is a Poisson random variable with mean n , and $n\tau_n$ is the expected vote profile, then the probability of a vote profile x in $Z(C)$ is

$$P(x | n\tau_n) = \prod_{c \in C} \left(\frac{\exp(n\tau_n(c)) (n\tau_n(c))^{x(c)}}{x(c)!} \right)$$

By the environmental-equivalence property of Poisson games (see Myerson 1998), any voter would also assign this probability to the event that x the profile of votes cast by all other voters.

The high scorer(s) in the vote profile x are

$$\Omega(x) = \operatorname{argmax}_{i \in \{1,2,3\}} \sum_{c \in C} x(c) c_i.$$

Assuming random tie-breaking, the probability that candidate i wins with the vote profile x is

$$W_i(x) = 0 \text{ if } i \notin \Omega(x), \text{ and } W_i(x) = 1/\#\Omega(x) \text{ if } i \in \Omega(x)$$

(where $\#\Omega(x)$ is the number of high scorers at x). Then the expected utility for a type- t voter is

$$U(x,t) = \sum_{i \in \{1,2,3\}} W_i(x) t_i$$

when the vote profile is x . (Recall that t_i here is the payoff to a type- t voter when i wins.)

Now let $x+[c]$ denote the vote profile that differs from x in that there is one additional c -ballot. With beliefs τ_n in a voting game of expected size n , voter of type t should want to choose a ballot c that maximizes the expected payoff

$$\sum_{x \in Z(C)} P(x | n\tau_n) U(x+[c],t).$$

A strategy function is a function $\sigma: T \rightarrow \Delta(C)$, where $\sigma(c|t)$ denotes the probability that any given type- t voter would cast the ballot c . So we may say that a strategy function σ is a best response to an expected vote profile $n\tau_n$ iff, for each type t and each feasible ballot c ,

if $\sigma(c|t) > 0$ then $c \in \operatorname{argmax}_{d \in C} \sum_{x \in Z(C)} P(x|n\tau_n) U(x+[d],t)$.

If everyone applied the strategy function σ , then the probability of a randomly sampled voter casting the ballot c would be

$$\sum_{t \in T} r(t) \sigma(c|t).$$

In this paper, we study strategy functions that would be optimal, even in very large games, and even when there is some arbitrarily small probability of voters choosing any given ballot by mistake. So let us say that a perfect asymptotic equilibrium is a strategy function σ together with a probability distribution η on the set of candidates such that, for each $\varepsilon > 0$, we can find an infinite sequence of beliefs $\{\tau_n\}$, indexed on $n \rightarrow \infty$, such that:

σ is a best response to the expected vote profile $n\tau_n$, $\forall n$,

$\lim_{n \rightarrow \infty} \tau_n(c) > 0$, $\forall c \in C$, and

$$\left| \lim_{n \rightarrow \infty} \tau_n(c) - \sum_{t \in T} r(t) \sigma(c|t) \right| \leq \varepsilon, \quad \forall c \in C.$$

$$\left| \eta_i - \lim_{n \rightarrow \infty} \sum_{x \in Z(C)} P(x|n\tau_n) W_i(x) \right| \leq \varepsilon, \quad \forall i.$$

We shall use this concept as our basic solution throughout this paper, so the term "equilibrium" here may be taken to mean such a perfect asymptotic equilibrium. The limiting win-probability η_i for each candidate i is included in this definition of equilibrium so that we will be able to distinguish two equilibria where voting behavior is the asymptotically the same but the winners are different, as in Example 4 below. (To show existence of asymptotic equilibria, we can apply an equilibrium existence theorem for finite Poisson games, such as Theorem 0 of Myerson 1999, to a sequence of large games that have been perturbed by small probabilities of new types that are dedicated to each action; and then we can choose appropriately convergent subsequences as $n \rightarrow \infty$, using the fact that there are only finitely many possible sets of best-response functions.)

We are assuming here (until Example 4 in Section 7) that voters' utility payoffs depend on their votes only through their effect on the winner of the election. Thus, in the calculus of rational voting, each voter must ask what is the probability that his vote could change the winner of the election from one candidate to another. So to compute optimal votes for any type of voter, we need to compute the probability of various pivot events in which one vote can change the winner of the election. Let $\Lambda(c,i,j)$ denote the event that one more c -ballot could change the winner from candidate i to candidate j ,

$$\Lambda(c,i,j) = \{x \in Z(C) \mid W_i(x) > W_i(x+[c]), W_j(x) < W_j(x+[c])\}.$$

Then let $\Lambda^*(i,j)$ denote the event that candidates i and j are in a close race where one vote could make a difference between one of them winning or the other,

$$\Lambda^*(i,j) = \cup_{c \in C} (\Lambda(c,i,j) \cup \Lambda(c,j,i)).$$

Let Λ^{**} denote the event that there is a close race between some pair of candidates,

$$\Lambda^{**} = \cup_{(i,j)} \Lambda^*(i,j).$$

As noted in Section 1, there is an obvious symmetry among the three candidates in the Condorcet cycle (in our Poisson-replicated version as well as the simple three-voter version), and any anonymous neutral voting rule will give us a voting game in which the three candidates are treated symmetrically. We would like to characterize the voting rules that generate multiple nonsymmetric equilibria for this social-choice situation. Unfortunately, analyzing whether such nonsymmetric equilibria exist in great generality can be difficult, so we simplify the question here by asking instead whether a particularly extreme form of nonsymmetry exists: a nonsymmetry in which one candidate essentially vanishes from the rational voting calculus.

Obviously a voter's decision only matters in the event that there is a close race, when all

other votes are counted. That is,

$$\operatorname{argmax}_{d \in C} \sum_{x \in Z(C)} P(x | n\tau_n) U(x+[d], t) = \operatorname{argmax}_{d \in C} \sum_{x \in \Lambda^{**}} P(x | n\tau_n) U(x+[d], t).$$

So we may say that a race between a pair of candidates i and j is serious if, conditional on the event that there is some close race, there is a positive probability of the close race involving i and j . More formally, let us say that the $\{i, j\}$ race is serious in an equilibrium σ iff σ can satisfy the definition of a perfect asymptotic equilibria with sequences $\{\tau_n\}$ that also satisfy the additional property

$$\lim_{n \rightarrow \infty} P(\Lambda^*(i, j) | n\tau_n) / P(\Lambda^{**} | n\tau_n) > 0.$$

Let us say that a candidate i is out of contention in an equilibrium σ iff, for every other candidate j , the $\{i, j\}$ race is not serious in the equilibrium σ . Finally, let us say that an equilibrium σ is discriminatory iff some candidate is out of contention.

3. Magnitudes of cones events in large Poisson games

We now develop some basic techniques for evaluating probabilities of events in large Poisson games, reviewing and extending results from Myerson (1999). In this section, we take as given the finite set of possible actions C , and we consider the limits of event probabilities for sequences of expected vote profiles $n\tau_n$ as $n \rightarrow \infty$.

Our first proposition is a direct consequence of Lemma 1 and Theorem 2 of Myerson (1999), where the function ψ is defined as

$$\psi(\theta) = \theta(1 - \log(\theta)) - 1, \quad \forall \theta > 0, \quad \text{and} \quad \psi(0) = -1.$$

(Here the log is natural logarithm, base e .) So ψ is a strictly concave continuous function with

$$\psi'(\theta) = -\log(\theta), \quad \text{and} \quad \max_{\theta \geq 0} \psi(\theta) = \psi(1) = 0.$$

Proposition 1. Suppose that $\{x_n\}$ is a sequence in $Z(C)$, and

$$\lim_{n \rightarrow \infty} x_n(c)/(n\tau_n(c)) = \alpha(c) > 0 \text{ and } \lim_{n \rightarrow \infty} \tau_n(c) = \tau(c) > 0, \forall c \in C.$$

Then $\lim_{n \rightarrow \infty} \log(P(x_n | n\tau_n))/n = \sum_{c \in C} \tau(c) \psi(\alpha(c))$.

Furthermore, for any vector $w = (w(c))_{c \in C}$ such that each $w(c)$ is an integer (possibly negative),

$$(3.1) \quad \lim_{n \rightarrow \infty} P(x_n - w | n\tau_n)/P(x_n | n\tau_n) = \prod_{c \in C} \alpha(c)^{w(c)}.$$

In this result, the critical quantity $x_n(c)/(n\tau_n(c))$ may be called the offset ratio for c at the vote profile x . This offset ratio describes the number of c -ballots in x_n as a fraction of the expected number of c -ballots.

For any set $R \subseteq \mathbb{R}^C$, the probability of the event R with expected vote profile $n\tau_n$ is

$$P(R | n\tau_n) = \sum_{x \in R \cap Z(C)} P(x | n\tau_n).$$

Given a sequence of probability distributions $\{\tau_n\}$ in $\Delta(C)$, we define the magnitude of such an event R to be

$$\mu(R) = \lim_{n \rightarrow \infty} \log(P(R | n\tau_n))/n.$$

Notice that the magnitude of any event must be zero or negative. If the magnitude $\mu(S)$ is a negative number, then the probability of S goes to zero exponentially like $e^{n\mu(S)}$ as $n \rightarrow \infty$. If the magnitude of an event R is strictly less than the magnitude of S , then the probability of R becomes infinitesimal relative to the probability of S . That is,

$$\text{if } \mu(R) < \mu(S) \text{ then } \lim_{n \rightarrow \infty} P(R | n\tau_n)/P(S | n\tau_n) = 0.$$

General results for evaluating the magnitudes of many events have been developed by Myerson (1999), but here we develop some slightly stronger results for events that have the structure of a cone in \mathbb{R}^C . In voting games, we are often interested in events that defined by

homogeneous linear inequalities, and so have the geometrical structure of a cone. Such events include, for example, the event that candidate 3 gets the highest score, and the event that candidates 1 and 2 are tied.

Suppose that K is a finite set, and $b_k(c)$ is a rational number for each k in K and each c in C . Let S denote a cone in \mathbb{R}^C defined by the conditions

$$(3.2) \quad S = \{x \in \mathbb{R}^C \mid x(c) \geq 0 \ \forall c \in C, \text{ and } \sum_{c \in C} b_k(c) x(c) \geq 0 \ \forall k \in K\}$$

Suppose also that there exists some vector x such that

$$x \in S, \text{ and } x(c) > 0 \ \forall c \in C,$$

so that the cone S is compatible with having large numbers of all possible ballots in C . Notice that a cone always includes the zero vector, and the event of zero turnout has magnitude -1 , because the probability of zero population is e^{-n} when population is a Poisson random variable with mean n . So the magnitude of a cone S cannot be less than -1 .

For any τ in $\Delta(C)$, consider the following magnitude problem on the cone S with beliefs τ :

$$(3.3) \quad \begin{aligned} &\text{maximize}_{\alpha} \sum_{c \in C} \tau(c) \psi(\alpha(c)) \text{ subject to} \\ &\sum_{c \in C} b_k(c) \tau(c) \alpha(c) \geq 0, \ \forall k \in K, \\ &\text{and } \alpha(c) \geq 0, \ \forall c \in C. \end{aligned}$$

Proposition 2 tells us that the solutions to this optimization problem are well behaved.

Proposition 2. For any τ in $\Delta(C)$, the magnitude problem on the cone S with beliefs τ has an optimal solution $\alpha = (\alpha(c))_{c \in C}$ in \mathbb{R}^C . The optimal value of the object function in this magnitude problem is a continuous function of τ in $\Delta(C)$. For each c in C , the value of $\alpha(c)$ in the optimal solution is unique and is a continuous function of τ in the domain where $\tau(c) > 0$.

The next proposition tells us that, at least when the limiting fractions for all ballots are positive, the magnitude of the cone S is the limiting solution to this optimization problem, and almost all probability in the cone is concentrated in the region where offset ratios are close to its unique solution.

Proposition 3. Suppose $\lim_{n \rightarrow \infty} \tau_n(c) = \tau(c) > 0$ for all c in C . Let α be an optimal solution of the magnitude problem on S with τ . Then

$$\alpha(c) > 0, \quad \forall c \in C,$$

$$\lim_{n \rightarrow \infty} \log(P(S | n\tau_n)) / n = \sum_{c \in C} \tau(c) \psi(\alpha(c)), \quad \text{and}$$

$$\lim_{n \rightarrow \infty} P(\{x \in S \mid \alpha(c) - \varepsilon \leq x(c) / (n\tau_n(c)) \leq \alpha(c) + \varepsilon, \forall c \in C\} \mid n\tau_n) / P(S \mid n\tau_n) = 1 \quad \forall \varepsilon > 0.$$

Furthermore, for any vector w such that $w(c)$ is an integer $\forall c$,

$$(3.4) \quad \lim_{n \rightarrow \infty} P(S - w \mid n\tau_n) / P(S \mid n\tau_n) = \prod_{c \in C} \alpha(c)^{w(c)}.$$

(Here $S - w = \{x - w \mid x \in S\}$.)

The cone structure of S enables us to apply duality to the magnitude problem. We may define dual of the magnitude problem on the cone S with beliefs τ to be the optimization problem:

$$(3.5) \quad \text{minimize}_{\lambda} \sum_{c \in C} \tau(c) \exp(\sum_k \lambda_k b_k(c)) - 1 \quad \text{subject to} \quad \lambda_k \geq 0 \quad \forall k \in K.$$

Our next proposition shows how the dual gives us a solution to the primal magnitude problem.

Proposition 4. Given any τ in $\Delta(C)$, let $D(\tau) = \{c \in C \mid \tau(c) > 0\}$. Suppose that there exists some x in the cone S such that $\{c \mid x(c) > 0\} = D(\tau)$. Then a vector α in \mathbb{R}^C is an optimal solution of the magnitude problem on S with τ if and only if the corresponding dual problem has an

optimal solution λ such that,

$$(3.6) \quad \alpha(c) = \exp\left(\sum_{k \in K} \lambda_k b_k(c)\right), \quad \forall c \in D(\tau).$$

Furthermore, the optimal value of the objective function in the magnitude problem on S with τ is equal to the optimal value of the objective function in the corresponding dual problem.

The proofs of these propositions are given in the Section 8 below.

4. Example 1: the Poisson-replicated Condorcet cycle

Now we can analyze our primary example, the large Poisson version of the Condorcet cycle. In this example, there are three candidates, numbered 1 to 3. There are three types of voters, and each type t is characterized by its utility vector $t = (t_1, t_2, t_3)$, where t_i denotes the utility payoff that the type- t voter would get if candidate i won the election. With this notation, voters in this example have three possible utility types: $(10, 5, 0)$, $(0, 10, 5)$, and $(5, 0, 10)$, yielding preferences $1 \succ 2 \succ 3$, $2 \succ 3 \succ 1$, and $3 \succ 1 \succ 2$ respectively. The classical version of this example had one voter of each type, as described in Section 1, but here we will consider a replicated version in which the number of voters is random and expected to be large. In our replicated version of the Condorcet cycle example, we assume that the number of voters is a Poisson random variable with some large mean denoted by n , and we suppose that each voter has an independent type that is equally likely to be one of these three types. That is, letting $r(t)$ denote the probability that a randomly sampled voter will have type t , this social choice situation is characterized by the distribution

$$r(10, 5, 0) = r(0, 10, 5) = r(5, 0, 10) = 1/3.$$

Then the numbers of voters of each type are independent Poisson random variables, each with mean $n/3$.

In this symmetric social choice situation, if there exist any nonsymmetric discriminatory equilibria then there exist discriminatory equilibria in which candidate 3 (say) is out of contention. So let us now begin to characterize the scoring rules for which there exists a discriminatory equilibrium of the Condorcet cycle example, where candidate 3 is out of contention.

In a discriminatory equilibrium where the only serious race is between candidates 1 and 2, every voter will want to cast a vote-vector that maximizes the difference between the points for his more-preferred among $\{1,2\}$ and his less-preferred among $\{1,2\}$. To prove this, notice first that all voters are assumed to have strict preferences among candidates 1 and 2. So let $L(1,2)$ denote the cone where candidates 1 and 2 are in a tie for first place. The pivot event $\Lambda^*(1,2)$ can then be characterized as a finite union of events of the form $L(1,2)-w$, where w is an integer vector in \mathbb{R}^C , chosen to make the scores of candidates 1 and 2 differ by one of the possible score-differences that has absolute value less than one. (With C containing finitely many rational vote vectors, the possible score differences is a discrete lattice in \mathbb{R} .) By the offset condition (3.4) in Proposition 3, each $\lim_{n \rightarrow \infty} P(L(1,2)-w | n\tau_n) / P(L(1,2) | n\tau_n)$ must be a strictly positive finite number, and so we also get

$$\begin{aligned} & \lim_{n \rightarrow \infty} P(L(1,2)-w | n\tau_n) / P(\Lambda^{**} | n\tau_n) \\ & = \lim_{n \rightarrow \infty} P(L(1,2)-w | n\tau_n) / P(\Lambda^*(1,2) | n\tau_n) > 0 \end{aligned}$$

(where the equality here uses the fact that $\{1,2\}$ is the only serious race). But suppose consider a voter who prefers candidate 1 over 2 (say), and consider any two vote-vectors d and \hat{d} such that $d_1 - d_2 > \hat{d}_1 - \hat{d}_2$. The event that this voter could make candidate 1 win instead of 2 by voting d

instead of \hat{d} includes $L(1,2)-[\hat{d}]$, and so the limiting conditional probability of this event given Λ^{**} must be positive, and so this event must in the limit be infinitely more likely than any pivot event involving candidate 3. So even if (say) the voter preferred candidate 3 most and \hat{d}_3 were greater than d_3 , the voter's expected payoff (against the expected vote profile $n\tau_n$, for all large n) would be increased by voting d instead of \hat{d} . Thus, the voter should choose a vote c that maximizes $c_1 - c_2$, because switching to such a vote would increase the probability of 1 winning instead of 2 by much more (proportionally) than any effect on candidate 3's probability of winning.

Without loss of generality, we are assuming that the ballots in the scoring rule have been normalized so that, in all feasible vote-vectors that maximize the difference between the highest and lowest candidates' points, the highest candidate gets 1 point and the lowest candidate gets 0 points. Scoring rules then may differ in the number of points that can be given to the middle-ranked candidate, in any feasible vote-vector that maximizes the difference of points between the other two candidates. Given any scoring rule, we let A denote the lowest number of points that can be given to candidate 3 in a feasible vote-vector that maximizes the difference of points between candidates 1 and 2; and we let B denote the highest number of points that can be given to candidate 3 in a feasible vote-vector that maximizes the difference of points between candidates 1 and 2. So when all voters have strict preferences among candidates 1 and 2, these two parameters (A,B) tell us everything that we need to know about the scoring rule for the purposes of characterizing discriminatory equilibria of voting games.

With the normalization assumed above, these parameters A and B must satisfy

$$0 \leq A \leq B \leq 1.$$

Geometrically, the set of pairs (A,B) that satisfy this constraint is a triangle with extreme points (0,0), (1,1), and (0,1) (as shown below in Figures 1 and 2). The extreme point (A,B) = (0,0) corresponds to plurality voting. Plurality voting and the scoring rules near it in this triangle may be called best-rewarding scoring rules. The extreme point (A,B) = (1,1) corresponds to negative-plurality voting. Negative-plurality voting and the scoring rules near it in this triangle may be called worst-punishing scoring rules. The extreme point (A,B) = (0,1) corresponds to approval voting, where voters can give either 0 points or 1 point to their middle candidate. Approval voting and the scoring rules near it in this triangle may be called flexible scoring rules. Borda voting corresponds to the point (A,B) = (1/2,1/2), the midpoint of the edge between plurality and negative-plurality voting.

When candidate 3 is out of contention in the Condorcet cycle example, the voters with the preference ordering $1 \succ 2 \succ 3$ will want to give candidate 3 the fewest possible points (A) consistent with maximally separating candidate 1 over candidate 2, and so they should all cast the vote-vector (1,0,A). The voters with the preference ordering $3 \succ 1 \succ 2$ will want to give candidate 3 the most possible points (B) consistent with maximally separating candidate 1 over candidate 2, and so they should all cast the vote-vector (1,0,B). Finally, the voters with the preference ordering $2 \succ 3 \succ 1$ should cast the vote-vectors (0,1,A) or (0,1,B) to maximally separate candidate 2 over candidate 1. So for some number ρ such that $0 \leq \rho \leq 1$, a discriminatory equilibrium of the Condorcet cycle where 3 is out of contention would have to be $\bar{\sigma}$ of the form

$$\begin{aligned}\bar{\sigma}((1,0,A)|(10,5,0)) &= 1, & \bar{\sigma}((1,0,B)|(5,0,10)) &= 1, \\ \bar{\sigma}((0,1,A)|(0,10,5)) &= \rho, & \bar{\sigma}((0,1,B)|(0,10,5)) &= 1-\rho.\end{aligned}$$

Let $\bar{\tau}$ be the expected vote profile that is generated by this strategy function $\bar{\sigma}$:

$$\bar{\tau}(1,0,A) = 1/3, \bar{\tau}(1,0,B) = 1/3, \bar{\tau}(0,1,A) = \rho/3, \bar{\tau}(0,1,B) = (1-\rho)/3.$$

The expected point scores with a expected population size n are then

$$2n/3, n/3, A(1+\rho)n/3 + B(2-\rho)n/3$$

for candidates 1, 2, and 3 respectively.

The highest magnitude in Λ^{**} must be achieved in a pivot event that involves the winner at the expected vote profile, because magnitude function is continuous and has a global maximum at the expected vote profile. So if candidate 3 were the winner at the expected vote profile, then candidate 3 could not be out of contention. Thus, a necessary condition for a discriminatory equilibrium is that

$$(4.1) \quad A(1+\rho) + B(2-\rho) < 2,$$

and candidate 1 must be the likely winner in any discriminatory equilibrium where 3 is out of contention.

The magnitude of a {1,2} tie at a limiting distribution τ is

$$\max_{\alpha} \sum_{c \in C} \tau(c) \psi(\alpha(c)) \text{ subject to } \sum_{c \in C} (c_2 - c_1) \tau(c) \alpha(c) = 0.$$

When we expect 2/3 of the voters to vote $c_2 - c_1 = -1$ and 1/3 to vote $c_2 - c_1 = 1$, this magnitude is, by duality,

$$\min_{\lambda} (2/3)\exp(-\lambda) + (1/3)\exp(\lambda) - 1 = (2/3)2^{-0.5} + (1/3)2^{0.5} - 1 = -0.0572.$$

That is, the optimal dual solution has $\lambda = 0.5 \log(2)$. The maximal magnitude in the {1,2}-tie event is achieved at a vote profile where votes for 1 are decreased from the expected profile by a multiplicative offset factor of $2^{-0.5}$, and votes for 2 are increased from the expected profile by an offset factor of $2^{0.5}$, which gives candidates 1 and 2 each a score of $2^{0.5}n/3$. The score for candidate 3 with these offset ratios from $n\bar{\tau}$ is

$$A(2^{-0.5} + \rho 2^{0.5})n/3 + B(2^{-0.5} + (1-\rho)2^{0.5})n/3$$

which is less than $2^{0.5}n/3$ whenever the condition (4.1) above is satisfied, and so such a vote profile is in the $\{1,2\}$ -pivot event $\Lambda^*(1,2)$. Thus, by continuity, for any limiting distribution τ that is close to $\bar{\tau}$, the magnitude of the $\{1,2\}$ -pivot event must be close to -0.0572 .

Let $M(3)$ denote the event that candidate 3 is high scorer (winning or tied to win). Since 3 loses at the expected vote profile, the maximal magnitude in the cone $M(3)$ must be achieved at its boundary, where 3 is involved in a close race. Thus, the magnitude of a close race involving 3 is the same as the magnitude of the cone $M(3)$. To check for a discriminatory equilibrium, we need to check whether the magnitude of $M(3)$ is less than -0.0572 .

Consider first the case where $B < 0.5$. In this case, candidate 3 cannot win with any positive offset ratios from the expected vote distribution $\bar{\tau}$, because the average of candidate 1's points per voter and candidate 2's points per voter is always 0.5, while candidate 3 gets less than 0.5 points per voter in any ballot that has positive probability under $\bar{\tau}$. So for any value of ρ , the magnitude of $M(3)$ at $\bar{\tau}$ is -1 (the magnitude of zero turnout). Thus, for any positive vote distribution τ near $\bar{\tau}$, the magnitude of a close race involving 3 must be near -1 , and so must be less than -0.0572 . Thus we can certainly find a discriminatory equilibrium with some $\bar{\sigma}$ as above, when $B < 0.5$.

By a similar argument, if $B = 0.5$ and $A < 0.5$, then the magnitude of $M(3)$ at $\bar{\tau}$ must be less than $-1/3$, because candidate 3 cannot win with the votes that have positive probability under σ unless all the voters who would give A points to candidate 3, at that is expected to be at least $1/3$ of the voters. So for any positive τ near $\bar{\tau}$, the magnitude of a close race involving 3 must be less than -0.33 , which is less than -0.0572 . Thus we can find a discriminatory

equilibrium with some $\bar{\sigma}$ as above, when $B = 0.5$ and $A < 0.5$.

For the voters with utility type $(0,10,5)$, the optimal choice between the ballots $(0,1,A)$ and $(0,1,B)$ will depend on whether a close race involving 3 is more likely to be with candidate 1, in which case they prefer $(0,1,B)$, or with candidate 2, in which case they prefer $(0,1,A)$. Since candidate 1 is the likely winner with the expected vote distribution $\bar{\tau}$, a natural conjecture is that the $\{1,3\}$ -close race is more likely than the $\{2,3\}$ -close race, and so these voters should prefer to vote $(0,1,B)$; that is, ρ should be 0 in the definition of $\bar{\sigma}$. In fact this conjecture is correct. To prove it, we need a technical lemma about the dual of the magnitude problem on $M(3)$. With any expected distribution τ , this dual on $M(3)$ is

$$(4.2) \quad \text{minimize}_{\lambda} \sum_{c \in C} \tau(c) \exp(\lambda_1(c_3 - c_1) + \lambda_2(c_3 - c_2)) - 1.$$

Lemma. Suppose $B > 0.5$. If $\bar{\sigma}$ as above is a discriminatory equilibrium where candidate 3 is out of contention, then for any positive distribution τ sufficiently close to $\bar{\tau}$, an optimal solution λ for the dual problem (4.2) on $M(3)$ must satisfy $\lambda_1 > \lambda_2$.

Proof See Section 8 below.

With $\lambda_1 > \lambda_2 \geq 0$, we know that the maximal magnitude of $M(3)$ must be achieved where candidates 1 and 3 are in a close race, because λ_1 is the Lagrange multiplier for the constraint that candidate 3's score is not less than 1's score, and this Lagrange multiplier can be positive only if its constraint is binding. If there is not also a close $\{2,3\}$ -race where this maximal magnitude is achieved, then the magnitude of a close $\{2,3\}$ -race must be strictly less than the magnitude of a close $\{1,3\}$ -race in the limit, and so all voters with preferences $2 > 3 > 1$ must prefer to vote $(0,1,B)$

rather than $(0,1,A)$. So we are left with the case where the maximal magnitude in $M(3)$ is achieved where both constraints bind, at a three-way tie. So we know that, in the limit, almost all probability in $M(3)$ is concentrated in the region where offsets from the expected vote distribution are close to the offsets α that solve the magnitude problem for $M(3)$. In this region, we can set up a one-to-one correspondence between the vote profiles x such that

$$3 \in \Omega(x + [0,1,B]) \text{ but } 2 \in \Omega(x + [0,1,A])$$

and vote profiles y such that

$$3 \in \Omega(y + [0,1,B]) \text{ but } 1 \in \Omega(y + [0,1,A]),$$

such that corresponding points x and y satisfy

$$\sum_{c \in C} x(c) c_3 = \sum_{c \in C} y(c) c_3,$$

$$\sum_{c \in C} x(c) c_1 = 1 + \sum_{c \in C} y(c) c_2,$$

$$1 + \sum_{c \in C} x(c) c_2 = \sum_{c \in C} y(c) c_1.$$

To set up this correspondence, for any number δ in the lattice of possible score differences between candidates for vote profiles in $Z(C)$, find an integer vector w_δ that minimizes

$\sum_{c \in C} |w_\delta(c)|$ subject to the constraints

$$\sum_c w_\delta(c) c_3 = 0, \quad \sum_c w_\delta(c) c_2 = \delta + 1 = -\sum_c w_\delta(c) c_1.$$

(The neutrality of the scoring rule implies that such a vector w_δ can always be found.) Then associate a vector x as above with a vector $y = x - w_\delta$ where $\delta = \sum_c x(c) (c_2 - c_1)$. But then by the offset formulas (3.1) and (3.6)

$$\begin{aligned} P(y|n\tau_n)/P(x|n\tau_n) &\approx \prod_c \alpha(c)^{w_\delta(c)} \\ &= \exp(\sum_c (\lambda_1(c_3 - c_1) + \lambda_2(c_3 - c_2))w_\delta(c)) \\ &= \exp((\lambda_1 - \lambda_2)(1 + \sum_c x(c)(c_2 - c_1))) \end{aligned}$$

> 1 , because $\lambda_1 > \lambda_1$ and $2 \in \Omega(x+[0,1,A])$.

So we have a one-to-one correspondence between outcomes x where the type-(0,10,5) voters lose from voting (0,1,B) instead of (0,1,A) and more likely outcomes y where these voters would gain the same amount from voting (0,1,B) instead of (0,1,A). Thus, in expected value, these voters must prefer to vote (0,1,B). So a discriminatory equilibrium $\bar{\sigma}$ must have $\rho = 0$.

This result leaves us only to compute the magnitude of $M(3)$ for the expected vote distribution $\bar{\tau}$ such that

$$\bar{\tau}(1,0,A) = 1/3, \quad \bar{\tau}(1,0,B) = 1/3, \quad \bar{\tau}(0,1,B) = 1/3$$

and all other $\bar{\tau}(c) = 0$. We have already seen that the magnitude of $M(3)$ must be achieved at the boundary where candidate 3 is tied with 1, so the magnitude of a $\{1,3\}$ -tie is not less than the magnitude of $M(3)$. Furthermore, the magnitude of a $\{1,3\}$ -tie can be more than the magnitude of $M(3)$ only if the most likely $\{1,3\}$ -tie occurs where 2 has a higher score than either 1 or 3, in which case there must be a $\{1,2\}$ -tie that is closer to the expected vote profile (where 1 wins) and so the $\{1,2\}$ -pivot event must have higher magnitude than either a $\{1,3\}$ tie or $M(3)$. So we have a discriminatory equilibrium if and only if the event of a $\{1,3\}$ -tie has a smaller magnitude than the magnitude of a $\{1,2\}$ -tie, which we have seen to be $2^{1.5}/3 - 1 = -0.0572$.

Now the dual of the magnitude problem on the $\{1,3\}$ -tie event with a limit distribution τ is the one-dimensional problem

$$\text{minimize}_{\lambda} \sum_{c \in C} \tau(c) \exp(\lambda(c_3 - c_1)) - 1.$$

The optimality condition on the dual variable λ is

$$0 = \sum_{c \in C} \tau(c) (c_3 - c_1) \exp(\lambda(c_3 - c_1)),$$

When τ equals $\bar{\tau}$ as above, then, the optimal λ for this dual must satisfy

$$0 = (A-1) \exp(\lambda(A-1)) + (B-1) \exp(\lambda(B-1)) + B \exp(\lambda B).$$

So if the λ that satisfies this equation also gives us

$$(1/3) \exp(\lambda(A-1)) + (1/3) \exp(\lambda(B-1)) + (1/3) \exp(\lambda B) - 1 \leq 2^{1.5}/3 - 1$$

then we have a discriminatory equilibrium. It is straightforward to verify that such λ exist when the when the (A,B) pair is below a curve that goes through the points

$$(0, 0.649), (0.1, 0.621), (0.2, 0.593), (0.3, 0.563), (0.4, 0.532), (0.5, 0.500)$$

This curve is plotted by the thick line in Figure 1. In this figure, scoring rules are represented by points (A,B) in the triangle with corners (0,0), (0,1), and (1,1), where $0 \leq A \leq B \leq 1$. The scoring rules that yield discriminatory equilibria for the Condorcet cycle are in the lower portion of this triangle, from the corner (0,0) up to the thick line. The scoring rules that do not allow discriminatory equilibria for the Condorcet cycle are in the upper portion of the triangle, above the thick line.

[INSERT FIGURE 1 ABOUT HERE.]

Plurality voting, for example, has (A,B) = (0,0) and yields three discriminatory equilibria for the Condorcet cycle. If the voters focus on the equilibrium in which candidate 3 is out of contention, then the voters with preference orderings $1 > 2 > 3$ and $3 > 1 > 2$ voters should all vote (1,0,0), while voters with preferences $2 > 3 > 1$ should all vote (0,1,0), and so candidate 1 will almost surely win. But if the voters focus on the equilibrium in which candidate 2 is out of contention, then the voters with preference orderings $3 > 1 > 2$ and $2 > 3 > 1$ voters should all vote (0,0,1), while voters with preferences $1 > 2 > 3$ should all vote (1,0,0), and so candidate 3 will almost surely win. So focal manipulation can be critical under plurality voting.

Approval voting, with $(A,B) = (0,1)$, does not admit discriminatory equilibria for the Condorcet cycle. If everyone thought that candidate 3 was out of contention, then the $1 \succ 2 \succ 3$ voters would vote $(1,0,0)$, the $2 \succ 3 \succ 1$ voters would vote $(0,1,1)$, and $3 \succ 1 \succ 2$ voters would vote $(1,0,1)$, and so a close race between candidates 1 and 3 (each with expected score $2n/3$) would actually be much more likely than a close race between candidates 1 and 2 (who has expected score $n/3$), thus contradicting the assumption that candidate 3 was out of contention.

Negative plurality voting, with $(A,B) = (1,1)$, also does not admit discriminatory equilibria for the Condorcet cycle. If everyone thought that candidate 3 was out of contention, then everyone would vote either $(1,0,1)$ or $(0,1,1)$, with the result that candidate 3 would actually win the election.

Borda voting, with $A = B = 0.5$, is a difficult boundary case which we can best understand by instead considering examples with slightly higher and lower values of A and B . So let us consider instead the cases of $A = B = 0.6$ and $A = B = 0.4$.

First we consider the scoring rule where the middle candidate must be given 0.6 points, that is $(A,B) = (0.6,0.6)$. If the voters believed that candidate 3 was out of contention, then the voters with preference orderings $1 \succ 2 \succ 3$ and $3 \succ 1 \succ 2$ voters would vote $(1,0,0.6)$, and the $2 \succ 3 \succ 1$ voters would vote $(0,1,0.6)$, which gives the expected vote distribution

$$\bar{\tau}(1,0,0.6) = 2/3, \quad \bar{\tau}(0,1,0.6) = 1/3.$$

But when everybody is expected to vote either $(1,0,0.6)$ or $(0,1,0.6)$, candidate 2 wins if less than 40% vote $(1,0,0.6)$, candidate 3 wins if between 40% and 60% vote $(1,0,0.6)$, and candidate 1 wins if more than 60% vote $(1,0,0.6)$ (which is the expected result). Candidates 1 and 2 are close to tied only when about 50% vote $(1,0,0.6)$, but in this case they both would be far behind

candidate 3, unless all voter turnout has vanished. So the magnitude of a {1,2}-pivot event is

$$\mu(\Lambda^*(1,2)) = -1 < \mu(M(3)) = \mu(\Lambda^*(1,3)) = -0.010$$

as we approach this expected vote distribution $\bar{\tau}$. Thus, the assumption that candidate 3 would be out of contention leads to the contradictory conclusion that the pivot event $\Lambda^*(1,3)$ has a strictly greater magnitude than the pivot event $\Lambda^*(1,2)$ for all limiting distributions τ near $\bar{\tau}$, and so the scoring rule with $A = B = 0.6$ does not admit any discriminatory equilibria.

Now consider the scoring rule where the middle candidate must be given 0.4 points, that is $(A,B) = (0.4, 0.4)$. If the voters believed that candidate 3 was out of contention, then the voters with preference orderings $1 \succ 2 \succ 3$ and $3 \succ 1 \succ 2$ voters would vote $(1,0,0.4)$, and the $2 \succ 3 \succ 1$ voters would vote $(0,1,0.4)$, which gives the expected vote distribution

$$\bar{\tau}(1,0,0.4) = 2/3, \quad \bar{\tau}(0,1,0.4) = 1/3.$$

When everybody is expected to vote either $(1,0,0.4)$ or $(0,1,0.4)$, candidate 2 wins if less than 50% vote $(1,0,0.4)$, and candidate 1 wins if more than 50% vote $(1,0,0.4)$ (which is the expected result). Candidate 3 cannot win at all when only these two ballots are used, unless total voter turnout goes to zero, which is an event of magnitude -1 . Thus, the pivot events $\Lambda^*(1,3)$ and $\Lambda^*(2,3)$ have magnitudes

$$\mu(\Lambda^*(1,3)) = \mu(\Lambda^*(2,3)) = \mu(M(3)) = -1,$$

while the pivot event $\Lambda^*(1,2)$ has magnitude

$$\mu(\Lambda^*(1,2)) = -0.0572.$$

Thus, the assumption that candidate 3 would be out of contention leads to the conclusion that the pivot events $\Lambda^*(1,3)$ and $\Lambda^*(2,3)$ would indeed have strictly smaller magnitude than the pivot event $\Lambda(1,2)$, and so we have found a discriminatory equilibrium for the Condorcet cycle under

the scoring rule with $A = B = 0.4$.

In this discriminatory equilibrium with $A = B = 0.4$, the expected scores for candidates 1, 2, and 3 are respectively $2n/3$, $n/3$, and $0.4n$. So candidate 1 is likely to win, and candidate 3 is most likely to be in second place. But even though candidate 3 is expected to get more points than candidate 2, the pivot event $\Lambda^*(1,2)$ has a strictly higher magnitude than the pivot event $\Lambda^*(1,3)$, because $-0.0572 > -1$. This result contradicts the assumption of Myerson and Weber (1993) that pivot probabilities should be ranked according to the expected scores of the candidates. In the Myerson-Weber (1993) calculus, the pivot event $\Lambda^*(1,3)$ is assumed to have infinitely greater probability than any other pivot event when candidates 1 and 3 have respectively the highest and second-highest expected scores. But the results of the Poisson model in this case show the flaws of this assumption. When voters only use the vote-vectors $(1,0,0.4)$ and $(0,1,0.4)$, the expected second-place candidate 3 can draw even with the expected first-place candidate 1 only by a relatively large $(0,1,0.4)$ vote, which would then move candidate 2 up to or ahead of candidate 3. In essence, the Myerson-Weber assumption failed here because the relative supporters of the expected second-place candidate against candidate 1 are also even stronger supporters of the expected third-place candidate.

Because of the importance and difficulty of estimating equilibria of multicandidate voting games under different voting rules, Myerson and Weber (1993) tried to simplify the analysis of large voting games by some admittedly naive assumptions about how the expected vote distribution τ might determine the relative magnitudes of pivot probabilities. The Myerson-Weber methodology appears to be seriously inappropriate in this case of $A=B=0.4$, in comparison with our more sophisticated (but much more difficult) Poisson methodology of this paper. When we

look at the set of all scoring rules, however, the differences between the two methodologies seems less significant. According the Myerson-Weber (1993) methodology, discriminatory voting equilibria would exist for the Condorcet cycle when $2B+A \leq 1$, which would put the upper boundary of the discriminatory region below but not far from the line in Figure 1 that we generated with the Poisson model. Thus, the simpler Myerson-Weber analysis and the more sophisticated Poisson analysis agree in the general conclusion that the Condorcet cycle has multiple discriminatory equilibria only for scoring rules that are relatively close to plurality voting in this figure.

5. Example 2: One rotten apple

In the Condorcet cycle, discriminatory equilibria break an obvious symmetry among the candidates and allow a wide scope for heresthetic focal manipulation by political leaders. Thus, the Condorcet cycle is an example where the existence of discriminatory equilibria may be seen as undesirable or problematic. But there are other social choice situations where the existence of discriminatory equilibria seems more desirable, and nonexistence of discriminatory equilibria becomes problematic. In this section, we consider a simple example of such a situation.

In this Example 2, there are again three candidates, numbered 1, 2, and 3, but there are only two types of voters. The first type of voters, whom we may call type-1, have utility $t = (10,5,0)$, so their preference ordering over the candidates is $1 \succ 2 \succ 3$. The second type of voters, whom we may call type-2, have utility $t = (5,10,0)$, so their preference ordering over the candidates is $2 \succ 1 \succ 3$. The number of voters is a Poisson random variable with a large mean n , and each voter is equally likely to be either of these two types. That is, letting $r(t)$ denote the

probability that a randomly sampled voter will have utility function t over the set of candidates, this Example 2 can be characterized by the distribution

$$r(10,5,0) = r(5,10,0) = 1/2.$$

Then the numbers of each type of voter in this example are independent Poisson random variables, each with mean $n/2$.

In many social choice situations (like the Condorcet cycle), the meaning of the democratic ideal that social choices should implement majority preferences can be obscure. In this simple Example 2, however, the meaning of this ideal seems clear: The ideal democratic winner should be candidate 1 if there are more type-1 voters, and should be candidate 2 if there are more type-2 voters. This ideal outcome would certainly occur if candidates 1 and 2 were the only candidates on the ballot. So in an ideal democratic system, the outcome of the election should not be affected by the presence of candidate 3, who is disliked by all of the voters. But we will see that, under some voting rules, candidate 3 may serve as a "rotten apple" who can spoil the election.

An event has magnitude 0 in a Poisson game if the event can occur when the offset ratio for each type (that is, the number of players of each type divided its expected value) is close to 1. So in the ideal democratic scenario for this example, the pivot event $\Lambda^*(1,2)$ must have magnitude 0, because it can occur when the numbers of the two types of voters exactly equal their expected values (each $n/2$). On the other hand, candidate 3 cannot win in this ideal democratic scenario when there is any positive number of voters, and so the pivot events $\Lambda^*(1,3)$ and $\Lambda^*(2,3)$ must have magnitude -1 (the magnitude of no voter turnout). So the democratic ideal can be achieved in equilibrium here if and only if there exists a discriminatory equilibrium where candidate 3 is out of contention.

If the voters think that candidate 3 is out of contention, then they will all maximally separate candidates 1 and 2 on the ballots, and will give the despised candidate 3 the fewest points possible subject to the constraint that he is the middle candidate in the ballot. So the type-1 voters will vote $(1,0,A)$, and the type-2 voters will vote $(0,1,A)$ if 3 is out of contention. That is, a discriminatory equilibrium where 3 is out of contention must have the expected vote distribution

$$\tau(1,0,A) = 1/2 = \tau(0,1,A).$$

It is straightforward to verify that, with this expected vote distribution τ , the pivot events $\Lambda^*(1,3)$ and $\Lambda^*(2,3)$ have magnitudes that are less than the magnitude of $\Lambda^*(1,2)$ if $A < 1/2$, and so Example 2 has a discriminatory equilibrium where candidate 3 is out of contention when $A < 1/2$. But when $A > 1/2$, this expected vote distribution τ would make candidate 3 the expected winner, and so 3 could not be out of contention. Thus, the dotted line at $A = 1/2$ in Figure 1 marks the boundary between the scoring rules that have discriminatory equilibria where 3 is out of contention (left of the dotted line) and scoring rules that do not have such discriminatory equilibria (right of the dotted line).

The region in the top left section of Figure 1, above the thick line and left of the dotted line, represents the set of scoring rules that do not generate discriminatory equilibria for the Condorcet cycle, where discriminatory equilibria seem problematic, but allow discriminatory equilibrium in Example 2, where discriminatory equilibria are desirable. So these two examples suggest the desirability of scoring rules like approval voting ($A=0, B=1$) which give voters a wide range of choices about how many points to give their middle-ranked candidate while maximally separating the other two. In comparison, the scoring rules in the bottom left region of Figure 1 (including plurality voting in the $A=B=0$ corner) may have too many discriminatory equilibria, and

the scoring rules in the top right region of Figure 1 (including negative-plurality voting in the $A=B=1$ corner) may have too few discriminatory equilibria.

6. Example 3: Symmetric equilibria when a bloc has two similar candidates

In our analysis of the Condorcet cycle, we compared voting rules in terms of the existence or nonexistence of nonsymmetric discriminatory equilibria. For the Condorcet cycle, because all three candidates are symmetric, a symmetric equilibrium of the Condorcet cycle would obviously give each candidate an equal $1/3$ probability of winning under any anonymous neutral voting rule. But we can also consider examples where the symmetry only involves two out of three candidates. For such situations, Cox (1987,1990), Myerson and Weber (1993), and Myerson (1993) have found considerable differences among electoral systems. In this section we consider the simplified version of Cox's model that was formulated in the final section of Myerson (1993), to generate results that can be compared to Figure 1.

So our Example 3 is a three-candidate election in which candidates 2 and 3 are equally desirable from the perspective of all voters. In this example there are just two types of voters. One type, whom we may call type-1 voters, prefer candidates 1 over candidates 2 and 3, say according to the utility vector $t=(10,0,0)$. The other other type in this population, whom we may call type-2 voters, prefer candidates 2 and 3 equally over candidate 1, say according to the utility vector $t=(0,10,10)$. Assume that the number of voters is a Poisson random variable with a large mean n , and suppose that each voter has an independent probability Q of being a type-1 voter who prefers candidate 1. So the expected type distribution r is

$$r(10,0,0) = Q, \quad r(0,10,10) = 1-Q.$$

Thus we have a situation in which there are two similar candidates seeking to represent the bloc of type-2 voters. Under some electoral systems, the rivalry between these similar candidates could hurt this bloc's chances of getting representation, unless the voters reach some pre-election understanding about concentrating their support behind one of the two rivals. Under other electoral systems, the bloc could win even if they failed to coordinate and simply treated the similar candidates symmetrically. It is easy to characterize these two categories of electoral systems for the class of scoring rules represented in Figure 1.

So let us now suppose that the parameters A and B in Figure 1 (where $A < B$) completely characterize the scoring rule. That is, suppose that the permissible vote-vectors in the scoring rule are $(1,A,0)$, $(1,B,0)$, and all permutations of these two vectors. (This assumption was without loss of generality only when we were studying discriminatory equilibria.) In the unique symmetric (with respect to candidates 2 and 3) equilibrium of this simple example, type-1 voters all randomize between the vote-vectors $(1,A,0)$ and $(1,0,A)$ with equal probability, while the type-2 voters all randomize between the vote-vectors $(0,B,1)$ and $(0,1,B)$ with equal probability. The critical question then is whether the expected total point score for candidate 1, which is Qn , is greater or less than the expected score for candidates 2 and 3, which for each is

$$0.5(1-Q)n + 0.5(1-Q)nB + 0.5QnA .$$

These expected point totals are equal when Q equals the critical value Q^* such that

$$(1-Q^*) + B(1-Q^*) + AQ^* = 2Q^*, \text{ that is,}$$

$$Q^* = \frac{1+B}{3+B-A} .$$

This quantity Q^* is Cox's threshold of diversity for these simple scoring rules with three

candidates. (See Cox 1987, 1990, and Myerson, 1993.) When the expected fraction of voters who prefer candidate 1 over candidates 2 and 3 is greater than this threshold Q^* , we have symmetric equilibria in which the candidate 1 will win almost surely. On the other hand, when the expected fraction of voters who prefer candidate 1 over candidates 2 and 3 is less than Q^* , we have symmetric equilibria in which winner will almost surely be one of the similar candidates 2 and 3.

Lines where the Cox threshold Q^* is constant are shown in Figure 2. Of particular interest is the line where $Q^* = 0.5$, which includes approval voting at $(A,B)=(0,1)$ and Borda voting at $(A,B)=(0.5,0.5)$. For scoring rules below this line, we can construct examples with $Q^* < Q < 0.5$ such that the single candidate of an expected minority bloc will win almost surely when the majority bloc has two similar candidates and treats them symmetrically. On the other hand, for scoring rules above this line, we can construct examples with $0.5 < Q < Q^*$ such that the single candidate of an expected majority bloc will almost surely lose when the minority bloc has two symmetric candidates on the ballot. From this perspective, $Q^* = 0.5$ seems to be the ideal.

[INSERT FIGURE 2 ABOUT HERE]

7. Example 4: An election with a vote-share effect

The above analysis assumed that voters are motivated purely by an interest in the winner of the election. But future political events may also depend on the share of the votes that each candidate gets in the current election, and so voters' payoffs may depend on the distribution of votes itself, not only on the identity of the winner. In this section we review an example from Piketty (1995) to show that, although such vote-share effects may of course change the analysis,

they do not eliminate the importance of pivot events in rational voters' calculations.

So as Example 4, let us consider an example that is adapted from Piketty (1995) to the Poisson framework. In this example, the election is a national referendum on the question of whether to ratify an international economic treaty. The treaty will be ratified if the number voting "Yes" is greater than the number voting "No." Let us suppose for simplicity that ratifying the agreement would offer substantial benefits to all voters. But suppose also that the higher is the percentage of votes for the treaty in this referendum, the weaker will be the bargaining position of this nation's representatives in subsequent international negotiations on other agreements. Let γ denote the fraction of "Yes" votes in the election results. Then we may suppose that the utility payoff U to each voter from the referendum is

$$U = 1 - g(\gamma) \text{ if } \gamma > 0.5, \quad U = -g(\gamma) \text{ if } \gamma \leq 0.5.$$

Here $g(\gamma)$ represents the expected loss from weakening the nation's position in subsequent negotiations, in a utility scale where the benefits from the treaty equal 1. Suppose that g is an increasing continuously differentiable function, so

$$g'(\gamma) > 0 \text{ for all } \gamma \text{ between } 0 \text{ and } 1.$$

The number of voters is a Poisson random variable with a large mean n .

With $g'(0) > 0$, there always exists an equilibrium in which everyone votes "No." But we now show that there also exist asymptotic equilibria in which each voter independently randomizes between voting "Yes" and "No" with positive probability. (We ignore any possibility of abstention here.)

Let $\tau_n(Y)$ and $\tau_n(N) = 1 - \tau_n(Y)$ denote the expected vote shares for "Yes" and "No" in an election of expected size n , and let $\tau(Y)$ denote a limit of $\tau_n(Y)$ as $n \rightarrow \infty$. For any large n , each

voter should figure that, if he votes "Yes" instead of "No" in the referendum, his vote will increase the expected loss in subsequent negotiations by the approximate amount

$$g'(\tau(Y))/n.$$

The pivot event that the treaty would be ratified if he voted "Yes" but would fail if he voted "No" is the event that, when everyone else's vote has been counted, there will be either an exact tie or the treaty will be ahead by one vote. This tie or near-tie event has magnitude

$$\mu = \min_{\lambda} \tau(Y) \exp(\lambda) + (1-\tau(Y)) \exp(-\lambda) - 1 = 2\sqrt{\tau(Y)(1-\tau(Y))} - 1.$$

(See formula (5.1) in Myerson, 1999.) When $\tau(Y) \neq 1/2$, this magnitude μ is strictly negative, and so the pivot probability must be going to zero an exponential rate, as $e^{n\mu}$, which becomes infinitesimal relative to $g'(\tau(Y))/n$. Thus, if $\tau(Y)$ is not $1/2$ then the expected benefits from possibly causing the treaty to be ratified become infinitesimal relative to the expected loss from weakening the subsequent bargaining position, and so each voter would prefer to vote against the treaty.

But now consider the case where $\tau_n(Y)$ converges to $1/2$ as $n \rightarrow \infty$. Each voter should think that, when all other votes are counted, the number of votes for the treaty and the number of votes against the treaty are independent Poisson random variables with means $n\tau_n(Y)$ and $n(1-\tau_n(Y))$ respectively. In the region where offset ratios are close to 1 (near the expected values), large Poisson random variables can be approximated by the integer-rounding of Normal random variables that have variance equal to the mean (see Theorem 3 in Myerson 1999). So the net surplus of votes for the treaty (that is, the number of votes for the treaty minus the number of votes against the treaty) can be approximated by the integer rounding of a Normal random variable with mean

$$n\tau_n(Y) - n(1 - \tau_n(Y)) = n(2\tau_n(Y) - 1)$$

and with variance

$$n\tau_n(Y) + n(1 - \tau_n(Y)) = n.$$

For large n , the probability that this random variable is equal to 0 or 1 is approximately

$$\frac{2 \exp(-0.5 n(2\tau_n(Y) - 1)^2)}{\sqrt{2\pi n}}$$

When $\tau_n(Y)$ is exactly 0.5, this probability must be much larger than the expected loss $g'(0.5)/n$ for all sufficiently large n , because $1/n$ becomes infinitesimal relative to $1/\sqrt{n}$. So there is some interval around 0.5, which shrinks as $n \rightarrow \infty$, such that, when $\tau_n(Y)$ is in this interval, the expected loss from weakening the subsequent bargaining position is smaller than the expected benefits from possibly causing the treaty to be ratified, and so each voter would prefer to vote for the treaty.

The voters become willing to randomize between voting for or against the treaty at the boundaries of this interval, which can be estimated by the equation

$$\frac{2 \exp(-0.5 n(2\tau_n(Y) - 1)^2)}{\sqrt{2\pi n}} = \frac{g'(0.5)}{n}.$$

This equation holds when

$$\tau_n(Y) = 0.5 \pm \frac{0.5}{\sqrt{n}} \sqrt{\log\left(\frac{2n}{(g'(0.5))^2 \pi}\right)}.$$

So we find two randomized equilibria, one in which $\tau_n(Y)$ is slightly bigger than 0.5, and another in which $\tau_n(Y)$ is slightly less. The asymptotic equilibrium strategy in either case is $\sigma(Y) = \sigma(N) = 1/2$, but the limiting probability of the "Yes" side winning the election is very different in the two sequences.

In the equilibrium where $\tau_n(Y)$ is bigger than 0.5, the net surplus of votes for the treaty

has mean

$$n(2\tau_n(Y)-1) = \sqrt{n} \sqrt{\log\left(\frac{2n}{(g'(0.5))^2 \pi}\right)}$$

and has standard deviation \sqrt{n} , and so the probability of the treaty being ratified (having a positive net surplus of votes) goes to one as $n \rightarrow \infty$. In the case where $g(\gamma) = \gamma$ and n is 100 million, for example, these equilibrium formulas give us $\tau_n(Y) = 0.500212$, and the difference between the votes for and against the treaty has expected value 42,400 with standard deviation 10,000. This equilibrium also has a certain stability, in the sense that if the expected number of votes for the treaty were larger than this equilibrium amount then some voters would have an incentive to switch to voting against the treaty, and if the expected number of votes for the treaty were slightly less than this equilibrium amount then some voters would have an incentive to switch to voting for the treaty. So in this equilibrium the pivot probability creates an incentive to vote for the treaty that successfully counteracts the vote-share effect's incentive to vote against the treaty.

In the other randomized equilibrium where $\tau_n(Y)$ is slightly less than 0.5, the treaty fails with high probability when n is large. But this equilibrium is not stable in the above sense.

8. Proofs of the Propositions and the Lemma

Proposition 1. Suppose that $\{x_n\}$ is a sequence in $Z(C)$, and

$$\lim_{n \rightarrow \infty} x_n(c)/(n\tau_n(c)) = \alpha(c) > 0 \text{ and } \lim_{n \rightarrow \infty} \tau_n(c) = \tau(c) > 0, \forall c \in C.$$

Then $\lim_{n \rightarrow \infty} \log(P(x_n | n\tau_n))/n = \sum_{c \in C} \tau(c) \psi(\alpha(c))$.

Furthermore, for any vector $w = (w(c))_{c \in C}$ such that each $w(c)$ is an integer (possibly negative),

$$\lim_{n \rightarrow \infty} P(x_n - w | n\tau_n)/P(x_n | n\tau_n) = \prod_{c \in C} \alpha(c)^{w(c)}.$$

Proof. The first conclusion of Proposition 1 follows from Lemma 1 of Myerson (1999),

with the continuity of the ψ function. The second conclusion is a special case of Theorem 2 of Myerson (1999). It also can be easily derived from the basic Poisson formula, because

$$P(x_n - w | n\tau_n) / P(x_n | n\tau_n) = \prod_{c \in C} (n\tau_n(c))^{-w(c)} x_n(c)! / (x_n(c) - w(c))!$$

which approaches $\prod_{c \in C} (x_n(c) / (n\tau_n(c)))^{w(c)}$ as the $x_n(c)$ become large. Q.E.D.

Proposition 2. For any τ in $\Delta(C)$, the magnitude problem on the cone S with beliefs τ has an optimal solution $\alpha = (\alpha(c))_{c \in C}$ in \mathbb{R}^C . The optimal value of the object function in this magnitude problem is a continuous function of τ in $\Delta(C)$. For each c in C , the value of $\alpha(c)$ in the optimal solution is unique and is a continuous function of τ in the domain where $\tau(c) > 0$.

Proof. The objective function $\sum_{c \in C} \tau(c) \psi(\alpha(c))$ and the constraints $\sum_{c \in C} b_k(c) \tau(c) \alpha(c)$ obviously do not depend on $\alpha(c)$ when $\tau(c) = 0$. So let $D(\tau) = \{c \in C \mid \tau(c) > 0\}$. When we ignore the irrelevant components of α for c not in $D(\tau)$, we get the problem of choosing $(\alpha(c))_{c \in D(\tau)}$ to maximize

$$\sum_{c \in D(\tau)} \tau(c) \psi(\alpha(c)),$$

which is a strictly concave function of $(\alpha(c))_{c \in D(\tau)}$, subject to the linear constraints

$$\sum_{c \in D(\tau)} b_k(c) \tau(c) \alpha(c) \geq 0, \quad \forall k,$$

$$\alpha(c) \geq 0, \quad \forall c \in D(\tau).$$

Letting α be the zero vector satisfies all constraints and yields a magnitude of -1 . But if

$\sum_{c \in D(\tau)} \tau(c) \alpha(c) > e$ then $\sum_{c \in D(\tau)} \tau(c) \psi(\alpha(c)) < \psi(e) = -1$. So our continuous objective function

must achieve a maximum within the compact set satisfying $\sum_{c \in D(\tau)} \tau(c) \alpha(c) \leq e$ and the other

linear constraints of S . Strict concavity of the objective guarantees that this optimal solution

$(\alpha(c))_{c \in D(\tau)}$ must be unique. The solution can then be extended to $\alpha \in \mathbb{R}^C$ by letting $\alpha(c)$ be any

nonnegative number when $\tau(c)=0$.

Now suppose that $\{\tau_n\}$ is a sequence in $\Delta(C)$ that converges to τ . For each n , let α_n be an optimal solution to the magnitude problem on S with τ_n . Let $y_n(c) = \alpha_n(c) \tau_n(c)$. We know that

$$\sum_{c \in C} y_n(c) = \sum_{c \in C} \alpha_n(c) \tau_n(c) \leq e,$$

and so y_n sequence must have a subsequence that is convergent in the compact set of nonnegative vectors that sum to less than e . Let y denote the limit of this sequence.

By construction, if $\tau_n(c) = 0$ then $y_n(c) = 0$. We could have a sequence with $\tau_n(c) > 0$ converging to $\tau(c)=0$, but in this case the $y_n(c)$ must converge to $y(c)=0$, because the term

$$\begin{aligned} \tau_n(c) \psi(\alpha_n(c)) &= \tau_n(c) \psi(y_n(c)/\tau_n(c)) = \\ &= \tau_n(c)((y_n(c)/\tau_n(c))(1 - \log(y_n(c)/\tau_n(c))) - 1) \\ &= y_n(c)(1 - \log(y_n(c)/\tau_n(c))) - \tau_n(c) \end{aligned}$$

would take our objective function to $-\infty$ as $n \rightarrow \infty$ if $y_n(c)$ converged to a positive limit while $\tau_n(c)$ went to zero; but we know that the optimal value of the object is bounded below by -1 . So $y(c)$ is 0 when $\tau(c)$ is 0. Then let $\alpha(c) = y(c)/\tau(c)$ if $\tau(c) > 0$, and let $\alpha(c) = 1$ if $\tau(c) = 0$. This vector α is feasible for the magnitude problem on S with τ because

$$\sum_c b_k(c) \tau(c) \alpha(c) = \sum_c b_k(c) y(c) = \lim_{n \rightarrow \infty} \sum_c b_k(c) \tau_n(c) \alpha_n(c) \geq 0, \quad \forall k.$$

The value of the objective at this vector α is

$$\sum_c \tau(c) \psi(\alpha(c)) = \lim_{n \rightarrow \infty} \sum_c \tau(c) \psi(\alpha_n(c)) \geq \lim_{n \rightarrow \infty} \sum_c \tau_n(c) \psi(\alpha_n(c)),$$

where the last inequality admits a possibility that the negative numbers $\tau_n(c)\psi(\alpha_n(c))$ might converge to a nonzero limit when $\tau_n(c) \rightarrow 0$ (which is the one case where $\alpha_n(c)$ does not converge to $\alpha(c)$). Thus, the value of the objective in our maximization problem with τ is not less than the limit of the values of the objective with τ_n .

Now suppose that β is an optimal solution for the magnitude problem on S with τ .

Without loss of generality, we can also assume that $\tau_n(c) > 0$ for all n in the sequence for all c such that $\tau(c) > 0$. So let $\beta_n(c) = \beta(c) \tau(c) / \tau_n(c)$ when $\tau(c) > 0$, and let $\beta_n(c) = \beta(c)$ if $\tau(c) = 0$.

Then β_n is feasible for the magnitude problem on S with τ_n , because

$$\sum_c b_k(c) \tau_n(c) \beta_n(c) = \sum_c b_k(c) \tau(c) \beta(c) \geq 0.$$

By optimality of α_n with τ_n , we get

$$\sum_c \tau_n(c) \psi(\alpha_n(c)) \geq \sum_c \tau_n(c) \psi(\beta_n(c)), \quad \forall n.$$

But β_n converges to β and τ_n converges to τ , and so by continuity

$$\lim_{n \rightarrow \infty} \sum_c \tau_n(c) \psi(\beta_n(c)) = \sum_c \tau(c) \psi(\beta(c)).$$

So we have

$$\sum_c \tau(c) \psi(\alpha(c)) \geq \lim_{n \rightarrow \infty} \sum_c \tau_n(c) \psi(\alpha_n(c)) \geq \sum_c \tau(c) \psi(\beta(c)),$$

which implies (by definition of β) that α is an optimal solution to the magnitude maximization problem on S with τ , and all three of these expressions must be equal. Thus the optimal value of the objective is a continuous function of the distribution τ in $\Delta(C)$, and the $\alpha(c)$ in the optimal solution depends continuously on the distribution τ when $\tau(c) > 0$. Q.E.D.

Proposition 3. Suppose $\lim_{n \rightarrow \infty} \tau_n(c) = \tau(c) > 0 \quad \forall c \in C$. Let α be an optimal solution of the magnitude problem on S with τ . Then

$$\alpha(c) > 0, \quad \forall c \in C,$$

$$\lim_{n \rightarrow \infty} \log(P(S | n\tau_n)) / n = \sum_{c \in C} \tau(c) \psi(\alpha(c)), \quad \text{and}$$

$$\lim_{n \rightarrow \infty} P(\{x \in S \mid \alpha(c) - \varepsilon \leq x(c) / (n\tau_n(c)) \leq \alpha(c) + \varepsilon, \forall c \in C\} | n\tau_n) / P(S | n\tau_n) = 1 \quad \forall \varepsilon > 0.$$

Furthermore, for any vector w such that $w(c)$ is an integer $\forall c$,

$$\lim_{n \rightarrow \infty} P(S-w | n\tau_n) / P(S | n\tau_n) = \prod_{c \in C} \alpha(c)^{w(c)}.$$

Proof. With all $\tau(c) > 0$, all $\alpha(c)$ must be nonzero at an optimal solution, because $\psi'(\theta)$ goes to $+\infty$ as θ goes to 0, and so for any vector x in S that has all positive components, the formula

$$\sum_{c \in C} \tau(c) \psi(\alpha(c) + \varepsilon x(c) / \tau(c))$$

would be an increasing function of ε when ε is close to 0 if some $\alpha(c)$ were equal to zero, while $(\alpha(c)\tau(c) + \varepsilon x(c))_{c \in C}$ would be in the cone S , which would contradict the optimality of α .

Theorem 1 of Myerson (1999) gives us

$$\begin{aligned} \lim_{n \rightarrow \infty} \log(P(S | n\tau_n)) / n &= \lim_{n \rightarrow \infty} \max_{x \in S} \sum_{c \in C} \tau_n(c) \psi(x(c) / n\tau_n(c)) \\ &= \lim_{n \rightarrow \infty} \max_{\alpha} \sum_{c \in C} \tau_n(c) \psi(\alpha(c)) \text{ subject to } (\alpha(c)\tau_n(c))_{c \in C} \in S. \end{aligned}$$

Then by Proposition 1 above

$$\begin{aligned} \lim_{n \rightarrow \infty} \max_{\alpha} \sum_{c \in C} \tau_n(c) \psi(\alpha(c)) \text{ subject to } (\alpha(c)\tau_n(c))_{c \in C} \in S \\ = \max_{\alpha} \sum_{c \in C} \tau(c) \psi(\alpha(c)) \text{ subject to } (\alpha(c)\tau(c))_{c \in C} \in S. \end{aligned}$$

Let $R = \{x \in S \mid \alpha(c) - \varepsilon \leq x(c) / (n\tau_n(c)) \leq \alpha(c) + \varepsilon, \forall c \in C\}$. The set $S \setminus R$ excludes a neighborhood of the unique optimum α for the magnitude problem on S , so

$$\begin{aligned} \lim_{n \rightarrow \infty} \max_{x \in S \setminus R} \sum_{c \in C} \tau_n(c) \psi(x(c) / (n\tau_n(c))) \\ < \sum_c \tau(c) \psi(\alpha(c)). \end{aligned}$$

So by Theorem 1 of Myerson (1999),

$$\lim_{n \rightarrow \infty} \log(P(S \setminus R | n\tau_n)) / n < \lim_{n \rightarrow \infty} \log(P(S | n\tau_n)) / n.$$

Thus $\lim_{n \rightarrow \infty} P(S \setminus R | n\tau_n) / P(S | n\tau_n) = 0$ and $\lim_{n \rightarrow \infty} P(R | n\tau_n) / P(S | n\tau_n) = 1$.

The final formula in Proposition 3 is an application of Theorem 2 in Myerson (1999) to the cone S . It also can be derived from the preceding result with the offset formula in

Proposition 1.

Q.E.D.

Proposition 4. Given any τ in $\Delta(C)$, let $D(\tau) = \{c \in C \mid \tau(c) > 0\}$. Suppose that there exists some x in the cone S such that $\{c \mid x(c) > 0\} = D(\tau)$. Then a vector α in \mathbb{R}^C is an optimal solution of the magnitude problem on S with τ if and only if the corresponding dual problem has an optimal solution λ such that,

$$\alpha(c) = \exp\left(\sum_{k \in K} \lambda_k b_k(c)\right), \quad \forall c \in D(\tau).$$

Furthermore, the optimal value of the objective function in the magnitude problem on S with τ is equal to the optimal value of the objective function in the corresponding dual problem.

Proof. Let x in S satisfy $\{c \mid x(c) > 0\} = D(\tau)$, and let β in \mathbb{R}^C satisfy $\beta(c)\tau(c) = x(c)$ for all c . Then as in the proof of Proposition 3, if $(\tau(c)\alpha(c))_{c \in C}$ is in S then $\alpha + \varepsilon\beta$ is a feasible offset vector for the magnitude problem on S with τ , for all $\varepsilon \geq 0$, and if $\alpha(c) = 0$ for some c in $D(\tau)$ then $\sum_{c \in C} \tau(c) \psi(\alpha(c) + \varepsilon\beta(c))$ must be strictly increasing for all sufficiently small ε (because $\psi'(0) = +\infty$). So an optimal solution of the magnitude problem on S with τ must have $\alpha(c) > 0$ for all c in $D(\tau)$. That is, the nonnegativity constraints are not binding at the optimal solution.

The Lagrangean associated with the magnitude problem on S with τ is

$$\sum_{c \in C} \tau(c) \psi(\alpha(c)) + \sum_{k \in K} \lambda_k \sum_{c \in C} b_k(c) \tau(c) \alpha(c)$$

The objective in this maximization problem is concave and differentiable, and the constraints are linear. So a necessary and sufficient condition for an optimum α is that there exist some vector of Lagrange multipliers $\lambda = (\lambda_k)_{k \in K}$ such that

$$0 = \tau(c) \psi'(\alpha(c)) + \sum_{k \in K} \lambda_k b_k(c) \tau(c), \quad \forall c \in C,$$

$$\lambda_k \geq 0 \quad \text{and} \quad \sum_{c \in C} b_k(c) \tau(c) \alpha(c) \geq 0, \quad \text{with at least one equality, } \forall k \in K.$$

But $\psi'(\alpha(c)) = -\log(\alpha(c))$. So the optimality equations give us equation (3.6):

$$\alpha(c) = \exp(\sum_{k \in K} \lambda_k b_k(c)), \quad \forall c \in D(\tau).$$

Then the remaining optimality conditions become, with $\lambda_k \geq 0$ for all k ,

$$\sum_{c \in C} b_j(c) \tau(c) \exp(\sum_{k \in K} \lambda_k b_k(c)) \geq 0, \text{ with equality when } \lambda_j > 0, \forall j \in K.$$

Now consider the dual problem

$$\text{minimize}_{\lambda} \sum_{c \in C} \tau(c) \exp(\sum_{k \in K} \lambda_k b_k(c)) - 1 \text{ subject to } \lambda_k \geq 0, \forall k \in K.$$

The derivative with respect to λ_j in this minimand is

$$\sum_{c \in C} \tau(c) b_j(c) \exp(\sum_{k \in K} \lambda_k b_k(c))$$

The objective in this problem is convex and differentiable, and so a necessary and sufficient condition for an optimal solution is that, for all j in K ,

$$\sum_{c \in C} \tau(c) b_j(c) \exp(\sum_{k \in K} \lambda_k b_k(c)) \geq 0, \text{ with equality if } \lambda_j > 0.$$

Thus, α is an optimal solution of the magnitude problem if and only if there exists some λ that is optimal for the dual, such that equation (3.6) is satisfied.

The optimal values of the magnitude problem and its dual are equal because, with equation (3.6) and the dual optimality conditions,

$$\begin{aligned} & \sum_{c \in C} \tau(c) \psi(\alpha(c)) \\ &= \sum_{c \in C} \tau(c) (\alpha(c)(1 - \log(\alpha(c))) - 1) \\ &= \sum_{c \in C} \tau(c) (\alpha(c) (1 - \sum_{k \in K} \lambda_k b_k(c)) - 1) \\ &= \sum_{c \in C} \tau(c) \alpha(c) - \sum_{k \in K} \lambda_k \sum_{c \in C} b_k(c) \tau(c) \alpha(c) - \sum_{c \in C} \tau(c) \\ &= \sum_{c \in C} \tau(c) \exp(\sum_{k \in K} \lambda_k b_k(c)) - 0 - 1. \end{aligned} \quad \text{Q.E.D.}$$

Lemma. Suppose $B > 0.5$. In the Condorcet cycle example, we consider the strategy

function $\bar{\sigma}$ such that $\bar{\sigma}((1,0,A)|(10,5,0)) = 1$, $\bar{\sigma}((1,0,B)|(5,0,10)) = 1$, $\bar{\sigma}((0,1,A)|(0,10,5)) = \rho$, and $\bar{\sigma}((0,1,B)|(0,10,5)) = 1 - \rho$, which generates the expected vote distribution $\bar{\tau}$ such that $\bar{\tau}(1,0,A) = 1/3$, $\bar{\tau}(1,0,B) = 1/3$, $\bar{\tau}(0,1,A) = \rho/3$, $\bar{\tau}(0,1,B) = (1 - \rho)/3$, and all other $\bar{\tau}(c) = 0$. If $\bar{\sigma}$ is a discriminatory equilibrium where candidate 3 is out of contention for the Condorcet cycle, then for any positive distribution τ sufficiently close to $\bar{\tau}$, an optimal solution λ for the dual problem on candidate 3's win-cone $M(3)$ must satisfy $\lambda_1 > \lambda_2$.

Proof. Let R denote the cone where twice candidate 3's score is greater than or equal to the sum of the other candidates' scores. The dual problem on this cone R may be written

$$\text{minimize}_\gamma G(\gamma, \tau) = \sum_{c \in C} \tau(c) \exp((2c_3 - c_1 - c_2)\gamma) - 1.$$

At an optimal solution of this problem, we have

$$\partial G(\gamma, \tau) / \partial \gamma = \sum_{c \in C} \tau(c) (2c_3 - c_1 - c_2) \exp((2c_3 - c_1 - c_2)\gamma) \geq 0 \text{ and } \gamma \geq 0,$$

and at least one of these inequalities must be satisfied as equality.

The vote distribution $\bar{\tau}$ assigns positive probability to $(1,0,B)$, in which candidate 3 gets more points than the average of candidates 1 and 2, given that $B > 0.5$, and so the hypothesis of Proposition 4 is satisfied at any τ near $\bar{\tau}$. So the optimal value of the dual is the same as the primal in a neighborhood of $\bar{\tau}$, and so Proposition 2 can be applied to show that the optimal γ is a continuous function of τ in a neighborhood of $\bar{\tau}$.

For any $\lambda = (\lambda_1, \lambda_2)$, let

$$F(\lambda, \tau) = \sum_c \tau(c) \exp(\lambda_1(c_3 - c_1) + \lambda_2(c_3 - c_2)) - 1.$$

Then the dual problem on the cone $M(3)$ with τ is

$$\text{minimize}_\lambda F(\lambda, \tau) \text{ subject to } \lambda_1 \geq 0 \text{ and } \lambda_2 \geq 0.$$

The derivatives of this dual objective are expected vote differences

$$\partial F(\lambda)/\partial \lambda_i = \sum_c (c_3 - c_i) \tau(c) \exp(\lambda_1(c_3 - c_1) + \lambda_2(c_3 - c_2)), \text{ for } i=1,2.$$

At the distribution $\bar{\tau}$, the dual objective is

$$\begin{aligned} F(\lambda, \bar{\tau}) &= (1/3) \exp(\lambda_1(A-1) + \lambda_2 A) + (1/3) \exp(\lambda_1(B-1) + \lambda_2 B) \\ &\quad + (\rho/3) \exp(\lambda_1 A + \lambda_2(A-1)) + ((1-\rho)/3) \exp(\lambda_1 B + \lambda_2(A-1)). \end{aligned}$$

Thus, at $\lambda = (\gamma, \gamma)$ (where γ is the optimal solution of the dual on \mathbb{R} with τ), we find

$$\begin{aligned} &\partial F((\gamma, \gamma), \tau)/\partial \lambda_1 + \partial F((\gamma, \gamma), \tau)/\partial \lambda_2 \\ &= \sum_{c \in C} \tau(c) (2 c_3 - c_1 - c_2) \exp((2 c_3 - c_1 - c_2)\gamma) \\ &= \partial G(\gamma, \tau)/\partial \gamma \geq 0, \text{ with equality if } \gamma > 0. \end{aligned}$$

But when $\tau = \bar{\tau}$, we also get

$$\partial F((\gamma, \gamma), \bar{\tau})/\partial \lambda_1 < \partial F((\gamma, \gamma), \bar{\tau})/\partial \lambda_2,$$

because

$$\begin{aligned} &\partial F((\gamma, \gamma), \bar{\tau})/\partial \lambda_1 - \partial F((\gamma, \gamma), \bar{\tau})/\partial \lambda_2 = \sum_c (c_2 - c_1) \bar{\tau}(c) \exp(\gamma(2c_3 - c_1 - c_2)) \\ &= (\rho/3 - 1/3) \exp(\gamma(2A-1)) + ((1-\rho)/3 - 1/3) \exp(\gamma(2B-1)) < 0. \end{aligned}$$

So for all τ sufficiently close to $\bar{\tau}$, we must have

$$\partial F((\gamma, \gamma), \tau)/\partial \lambda_2 > 0.$$

Notice also that (γ, γ) cannot be an optimal solution of the dual on $M(3)$. Even in the case where γ is 0 (where the above inequality does not prove suboptimality of (γ, γ) for the dual on $M(3)$), we cannot have an optimum where λ is $(0,0)$, because candidate 1 beats candidate 3 at the distribution $\bar{\tau}$, and so

$$\partial F((0,0), \tau)/\partial \lambda_1 = \sum_c (c_3 - c_1) \tau(c) < 0$$

for all τ close enough to $\bar{\tau}$.

Now let λ denote an optimal solution of the dual problem on $M(3)$ with τ . Then by convexity of F ,

$$\begin{aligned} F((\gamma, \gamma), \tau) &> F(\lambda, \tau) \\ &\geq F((\gamma, \gamma), \tau) + (\lambda_1 - \gamma) \partial F((\gamma, \gamma), \tau) / \partial \lambda_1 + (\lambda_2 - \gamma) \partial F((\gamma, \gamma), \tau) / \partial \lambda_2 \\ &= F((\gamma, \gamma), \tau) + (\lambda_2 - \lambda_1) \partial F((\gamma, \gamma), \tau) / \partial \lambda_2 + \lambda_1 \partial G(\gamma, \tau) / \partial \gamma. \end{aligned}$$

Thus, we must have $\lambda_2 < \lambda_1$.

Q.E.D.

REFERENCES

- G. Cox, "Electoral equilibrium under alternative voting institutions." American Journal of Political Science 31 (1987), 82-108.
- G. Cox, "Centripetal and centrifugal incentives in electoral systems." American Journal of Political Science 34 (1990), 903-935.
- E. Muller and M. Satterthwaite, "The equivalence of strong positive association and strategy-proofness," Journal of Economic Theory 14 (1977), 412-418.
- R. Myerson, "Incentives to cultivate favored minorities under alternative electoral systems." American Political Science Review 87 (1993), 856-869.
- R. Myerson, "Population uncertainty and Poisson games," International Journal of Game Theory 27 (1998), 375-392.
- R. Myerson, "Large Poisson games," Northwestern University Math Center discussion paper #1189 (1999).* To appear in Journal of Economic Theory.
- R. Myerson and R. Weber. 1993. "A theory of voting equilibria." American Political Science Review 87 (1993), 102-114.
- T. Piketty, "Voting as Communicating," CEPREMAP discussion paper (Paris, 1995).

*Adobe Acrobat PDF versions of these Northwestern discussion papers are available on the Internet at <http://www.kellogg.nwu.edu/research/math/nupapers.htm>

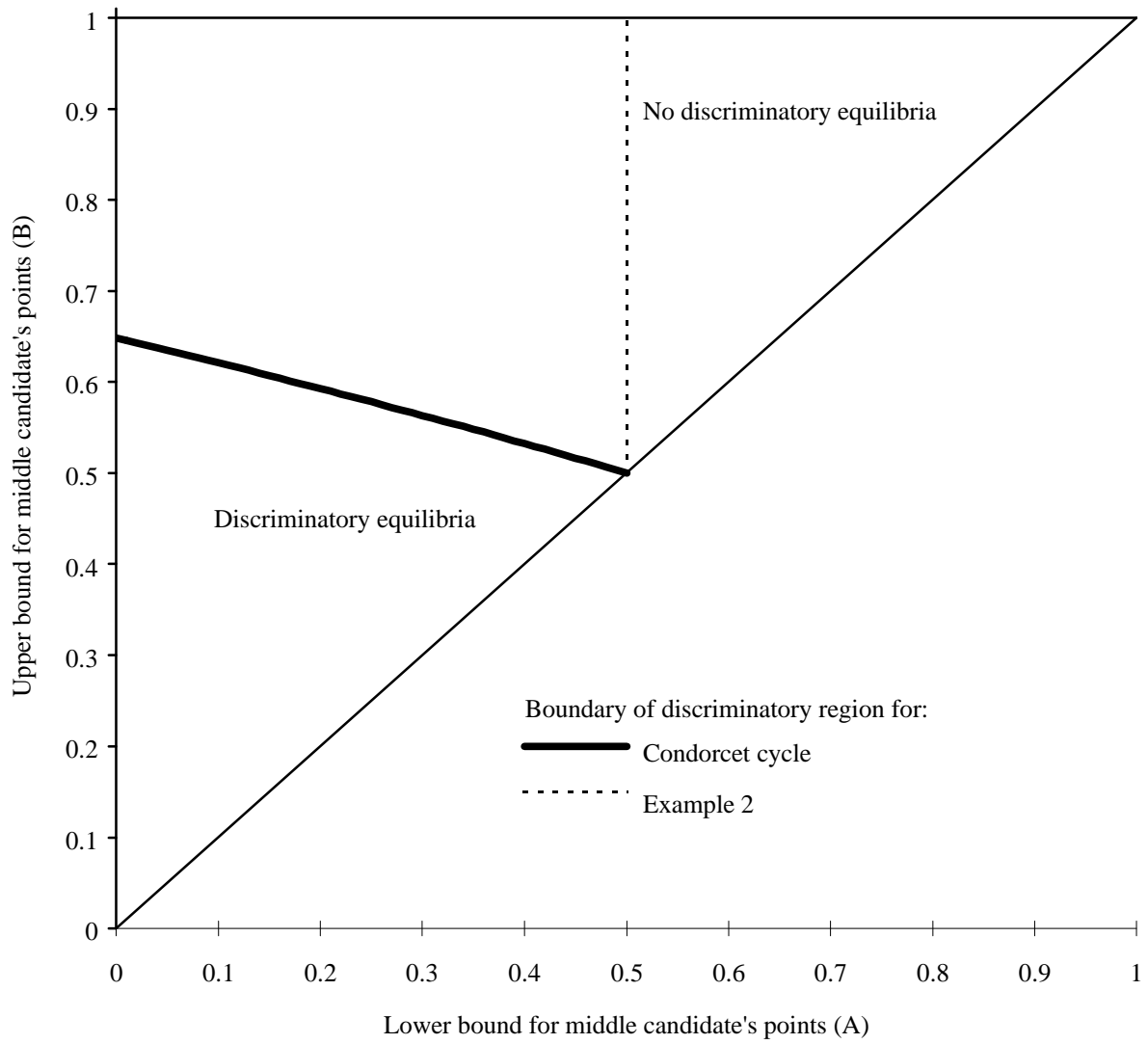


Figure 1. Existence of Discriminatory Equilibria.

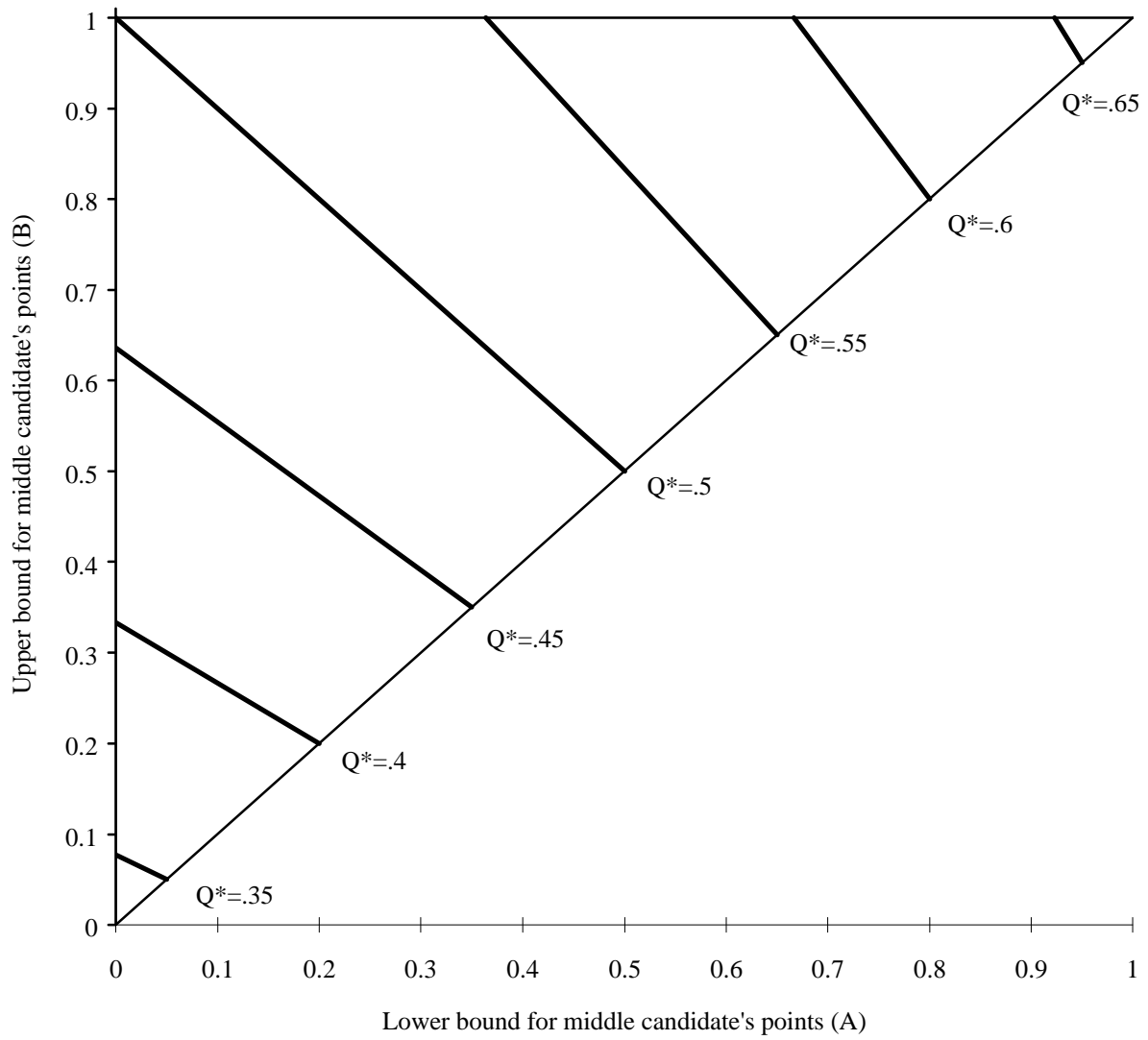


Figure 2. Level lines of Cox's threshold of diversity Q^* .