

Center for Mathematical Studies in Economics and Management Science
Northwestern University, Evanston, Illinois 60208

Discussion Paper No. 1000

INCENTIVES TO CULTIVATE FAVORED MINORITIES

UNDER ALTERNATIVE ELECTORAL SYSTEMS

by

Roger B. Myerson*

September 1992

Abstract. A simple model is used to compare, under different electoral systems, the incentives for candidates to create inequalities among otherwise homogeneous voters, by making campaign promises that favor small groups, rather than appealing equally to all voters. In this game model, each candidate generates offers for voters independently out of a distribution that is chosen by the candidate, subject only to the constraints that offers must be nonnegative and have mean 1. Symmetric equilibria with sincere voting are analyzed for two-candidate elections, and for multicandidate elections under rank-scoring rules, approval voting, and single transferable vote. Voting rules that can guarantee representation for minorities in multiseat elections generate, in this model, the most severely unequal campaign promises.

Acknowledgements. This research was initially inspired by conversations with Barry Ames and John Londregan. In writing this paper, I have also benefitted from many discussions with Robert Weber. Support from the Dispute Resolution Research Center is also gratefully acknowledged.

*Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208.

INCENTIVES TO CULTIVATE FAVORED MINORITIES

UNDER ALTERNATIVE ELECTORAL SYSTEMS

by Roger B. Myerson

Introduction

In an election campaign with many competing candidates, a candidate can try to appeal broadly to all voters equally, or a candidate may concentrate more narrowly on winning the support of minorities or special interest groups. Which strategy is better for winning an election? In this paper, we find that the answer to this question may depend crucially on the rules of the electoral system.

The focus here is on the creation of favored minority groups in election campaigns, so we consider situations in which all voters are initially the same, and we consider equilibria which are symmetric both across candidates and across voters. Nonetheless, we can show that some electoral systems may encourage candidates to build winning support by proposing transfer schemes for the benefit of a small minority of voters, thus breaking the symmetry among voters and creating favored minorities by their campaign promises. Under plurality voting, for example, if there are 10 candidates in the race then a candidate could easily win with only 20% of the vote, and so a campaign promise to tax 80% of the voters for the benefit of a 20% minority could be a successful strategy. We will show, however, that other electoral systems (such as Borda voting and approval voting) may significantly reduce such incentives to favor minority interests.

In general, there are several different reasons why a politician may want to make campaign promises that pander to special interest groups. A candidate

may make promises to a special group of voters in order to get votes from them, or to get donations for his campaign expenses, or to get bribes and kick-backs for his personal use. In this paper, we focus exclusively on first of these incentives. (The second requires a theory of how campaign spending influences votes; see Morton and Myerson 1992, for example. The third is essentially the topic of Myerson 1993.)

If a candidate could get away with it, he (or she) might try to win by promising everything to every voter. So we must assume that there are some limits on a candidate's ability to misrepresent himself as all things to all people. In this paper, we assume full rational expectations, so that each candidate must make promises that he can actually fulfill if elected. Thus, in our model, each candidate's campaign promises must satisfy a budget constraint, which is determined by the quantity of public resources that he would be able to allocate if he won the election.

For example, suppose that there are 1 million voters, and a winner of the election would have \$1 million of public resources to allocate among these voters. Then a candidate cannot make promises which average more than \$1 per voter. The equitable solution would be, of course, to promise \$1.00 to every voter. In a two-candidate race, such an equitable candidate would lose, however, if the second candidate promised \$1.25 to 80% of the voters and \$0 to the other 20% of the voters, which is also credible because the promises average to \$1 per voter. The first candidate could beat this second candidate, however, by switching to a strategy of promising \$1.50 to half of the voters and \$0.50 to the other half, but this strategy could also be beaten by other feasible strategies for the second candidate.

To identify optimal strategies in a Nash equilibrium, we must first

specify the technical details of the game, and we must keep the formulation simple enough to be tractable. In this paper, I consider one such simple formulation. To introduce the model in its simplest case, I begin by considering the two-candidate case. The analysis is then extended to cover all rank-scoring rules with any given number of candidates. The characterization of the unique symmetric equilibrium for general rank-scoring rules is the main result of this paper (Theorem 2). In subsequent sections, I also analyze the equilibrium offer distributions for approval voting and for single transferable vote, and I then extend these results to elections for multiseat councils.

The last section of the paper examines the close relationship between the results of this paper and Cox's (1987,1990) analysis of centripetal and centrifugal incentives in electoral systems. We find that electoral systems which encourage more diversity of candidates' positions, according to Cox's analysis, generally also incite candidates to create more inequality among voters, according to the results in this paper. That is, electoral systems which encourage candidates to advocate the interests of existing minorities may also incite candidates to use narrow campaign strategies that create favored minorities, even in situations where all voters are initially the same. These results should offer insights of practical importance for the constitutional design of democratic institutions.

Two-candidate elections

We begin by studying a race between two candidates, numbered 1 and 2, who are competing for some office. Each voter votes for one of the two candidates, and the winner is the candidate with the majority of the votes. (In a winner-take-all two-candidate election, this is essentially the only neutral

anonymous electoral system, so comparison of electoral systems will have to be deferred until we consider multicandidate elections, beginning in the next section.)

In our model, each candidate's strategy for generating campaign promises is described mathematically by an offer distribution, which is a probability distribution over the nonnegative real numbers. In this offer distribution, the probability of any interval may be interpreted as fraction of the voters for whom the candidate's offer of government benefits has a value that is in this interval. Equivalently, if a candidate's offer distribution has a cumulative probability function $F(\cdot)$, then $F(x)$ denotes the probability that a randomly sampled voter will be someone to whom the candidate has promised benefits that are worth less than x dollars.

Given that we want to investigate questions about the relative equality or inequality of the distribution of public benefits, the simplest possible way to formulate the government's service technology is to assume that benefits are derived purely from the allocation of linearly divisible resources. So each candidate's budget constraint is expressed, in our model, as a constraint on the average offer per voter that a candidate can promise. Specifically, we assume here that each candidate's offer distribution must have mean 1, to be considered credible by the voters.

To avoid specifying the number of voters and dealing with the complexities of large finite numbers, we will consider the number of voters to be essentially infinite. We should, however, interpret such an infinite model as an approximation to a large finite population with millions of voters. Under this interpretation, the budget constraint tells us that the winner of the election will control resources which have a monetary value that is equal in

magnitude to the number of voters. The winner will divide these resources and allocate them to the voters in whatever way he has promised during the campaign.

If candidate 2 could make offers after observing candidate 1's offers to all voters, then candidate 2 could always win the election. For example, candidate 2 could identify a small group of voters who are promised the most by candidate 1 (say, the top 10% in candidate 1's distribution) and offer nothing to this group. Then candidate 2 could offer every other voter slightly more than candidate 1 has promised him (or her), where the excess over candidate 1's offers is financed from the resources not given to the voters in the first group. So every voter outside of the first small (10%) group would vote for candidate 2, who would win by a landslide (90%). To avoid this simple outcome, we assume here that the two candidates must make their campaign promises independently.

We want to consider an electorate that is initially homogeneous, before the candidates make campaign promises. Thus, we assume that candidates' offers are independent across individual voters, so that no voter's offers have any specific relationship with any other set of voters' offers. That is, each voter's offer from each candidate is assumed to be drawn from the candidate's offer distribution independently of the candidate's offers to all other voters, and independently of all other candidates' offers to all voters. This independence assumption greatly simplifies our analysis, because it allows us to completely characterize a candidate's campaign promises by the marginal distribution of his individual offers to voters, without saying anything more about the joint distribution of offers to various sets of voters. The infinite-population assumption was introduced above essentially only to justify

this simplifying assumption of independence across voters.¹

We also assume that all offers must be nonnegative. One way to justify a nonnegativity assumption is to suppose that taxes are not in question in this campaign, and the candidates are only debating how the tax revenues should be redistributed to the voters. A more sophisticated model would allow taxes to be variable, but would recognize that tax-avoidance activities would increase the total economic cost of raising each additional dollar of revenue, as the tax rate increases. In such a framework, our model can be justified by supposing that there are no such tax-avoidance costs as long as taxes are below some maximal level, but tax-avoidance activities would make it infinitely costly to extract any additional tax revenue above this maximal level. To justify the 1-mean budget constraint, we can redefine the unit of money (if necessary) so that these maximal taxes would average one monetary unit per voter.

So at the beginning of the game, each candidate i simultaneously and independently chooses an offer distribution. We may represent candidate i 's offer distribution by a cumulative distribution function F_i , where $F_i(x)$ denotes the fraction of the voters to whom candidate i will offer less than x . Each offer distribution must have mean 1, and so F_i must be a nondecreasing function that satisfies

$$\int_0^{\infty} x \, dF_i(x) = 1,$$

as well as

$$F_i(x) = 0, \quad \forall x \leq 0,$$

and

$$\lim_{x \rightarrow \infty} F_i(x) = 1.$$

Next, each voter goes to each candidate's campaign headquarters and gets an

offer that is randomly drawn from the candidate's offer distribution, independently of all other offers. (We may suppose that the voter must show his voter's identification number when he gets his draw, to fix it verifiably and prevent him from coming back for a second draw if it is low.) Then, on election day, each voter knows how much each candidate has offered him, and he votes for the candidate who has promised him more. The winning candidate is the one who gets the most votes.

Let p_{ih} denote the probability that a randomly sampled voter would prefer candidate i over candidate h . So in Riemann-Stieltjes integral notation²,

$$p_{ih} = \int_0^\infty F_h(x) dF_i(x).$$

That is, the probability that any given voter prefers i to h is the expected value of $F_h(x)$ (the probability that h offers less than x to this voter) when x a random variable drawn from the F_i offer distribution. With a voting population that approaches infinity, the fraction of the voters who prefer i over h is almost surely p_{ih} . So candidate 1 wins (almost surely) if $p_{12} > p_{21}$, and candidate 2 wins if $p_{21} > p_{12}$. We assume throughout this paper that each candidate is motivated purely by the objective of winning the election. Thus, we can formulate the candidates' competition as a two-person zero-sum game in which the payoffs to candidates 1 and 2 respectively are $(1, -1)$ if $p_{12} > p_{21}$, $(-1, 1)$ if $p_{12} < p_{21}$, and $(0, 0)$ if p_{12} and p_{21} are equal.³

Because this game is symmetric, both candidates must get expected payoffs of 0 in equilibrium. We look for a symmetric equilibrium, in which both candidates use the same offer distribution.

Theorem 1. In a two-candidate race, the unique symmetric equilibrium is for each candidate to generate offers from a uniform distribution over the interval from 0 to 2. That is, the cumulative distribution F of the

equilibrium offer distribution for each candidate must satisfy

$$F(x) = x/2 \quad \text{if } 0 \leq x \leq 2.$$

This result is a special case of Theorem 2, which is proven in the next section, so we do not give a formal proof of uniqueness here. However, it is worthwhile to verify here that this uniform distribution on $[0,2]$ does give us an equilibrium. When both candidates draw their offers independently from the uniform distribution, they each expect to get half of the votes, of course. If candidate 1 stays with the uniform distribution on $[0,2]$ while candidate 2 deviates to some other distribution F_2 , then candidate 2's expected vote share is

$$\int_0^{\infty} F_1(x) dF_2(x) \leq \int_0^{\infty} (x/2) dF_2(x) = (\int_0^{\infty} x dF_2(x))/2 = 1/2$$

(because $F_1(x) = x/2$ when $0 \leq x \leq 2$, and $F_1(x) = 1 \leq x/2$ when $x > 2$, and the mean of candidate 2's distribution must equal 1). So candidate 2's vote share cannot be increased by deviating to another distribution.

This result is not really new. Game theorists since Gross and Wagner (1950) (see also Owen 1982, pages 78-83, and Shubik 1970) have found such uniform distributions in equilibria of "Colonel Blotto" games, where each of two competitors divides a fixed supply of resources over a set of battlefields. To apply the "Colonel Blotto" terminology to our context, we can reinterpret battlefields as voters, for whose votes the candidates compete with their offers. The hardest part of Gross and Wagner's problem was to construct joint distributions for allocations which always sum to the given total but which give uniform marginal distributions for each battlefield/voter. We have avoided such difficulties here by allowing the offers to be made independently to the various voters and by only requiring that the budget constraint be satisfied in expected value.

The advantage of our simplified formulation is that it will enable us to go beyond this "Colonel Blotto" literature and get results about more complicated situations in which more than two candidates are competing. Such games will allow us to get important new insights into the comparison of multicandidate electoral systems.

Multicandidate competition with rank-scoring rules

We consider now an election there are K candidates for a single indivisible office, and in which the voters vote according to a rank-scoring election rule. Such an electoral system can be characterized by an ordered sequence of K numbers, which we denote here by s_1, s_2, \dots, s_K , where

$$1 = s_1 \geq s_2 \geq \dots \geq s_K = 0.$$

Without any loss of generality, we normalize the highest and lowest point values to be 1 and 0 respectively. In the election, each voter must indicate a ranking of the K candidates on his ballot; then this voter's ballot gives the top-ranked candidate s_1 points, the second-ranked candidate s_2 points, and so on, with the j 'th ranked candidate getting s_j points. Each candidate's score is the average of the points that he gets from all voters. The winner of the election is the candidate with the highest score.

Plurality voting (or single nontransferable vote) is a rank-scoring rule in which $s_1 = 1$ and $s_j = 0$ for all $j > 1$. More generally, for any V between 1 and K , a rule where each voter must distribute V noncumulative votes on his ballot is represented by letting $s_1 = s_2 = \dots = s_V = 1$, and $s_j = 0$ for all $j > V$. Negative-plurality voting (in which each voter votes against one candidate and winner is the candidate with the fewest such votes against him) can be characterized as a rank-scoring rule in which $s_j = 1$ for all $j < K$, and

only $s_K = 0$. Borda rule is a rank-scoring rule in which $s_j = (K - j)/(K - 1)$, for each j .

Given any rank-scoring rule (s_1, s_2, \dots, s_K) , let \bar{s} denote the average of the ranking points that a voter can give, so

$$\bar{s} = \sum_{j=1}^K s_j / K.$$

Thus, $\bar{s} = V/K$ in a voting system where each voter must vote for exactly V of the K candidates, and $\bar{s} = 1/2$ for Borda rule.

We assume as before that, at the beginning of the campaign, each candidate must choose an offer distribution, which is a probability distribution over the nonnegative real numbers that has expected value equal to 1. During the campaign, each voter gets an offer or promised payoff from each candidate that is independently drawn from the candidate's offer distribution. Thus, a distribution in which most of the probability mass is near 1 would denote a campaign strategy of appealing broadly to most of the voters. On the other hand, a campaign strategy of offering special favors to small groups would be denoted by a distribution that has positive mass at numbers much higher than 1, balanced by a large mass near 0.

A symmetric equilibrium of this game is a scenario in which every candidate is expected to use the same offer distribution, and each candidate finds that using this offer distribution maximizes his chances of winning the election, when he knows that all other candidates are also independently allocating their offers according to this distribution, and all voters perceive that the K candidates all have the same probability of winning the election. In this paper, we focus exclusively on finding such symmetric equilibria. (The existence and significance of nonsymmetric equilibria of voting games has been examined by Myerson and Weber 1993, and Myerson 1993.)

As a consequence of the symmetry assumption, we can suppose here that the voters vote sincerely. Each voter thinks that, in the event that his vote could actually decide a close race between two of the candidates, it is equally likely to be any pair of the K candidates who are in this close race. (In the terms of Myerson and Weber 1993, we are assuming that each voter perceives that all of the $K(K-1)/2$ pivot probabilities are equal.) Thus, each voter should rank the candidates in order of their offers to him, giving s_1 points to the candidate who offers him the most, s_2 points to the candidate who offers him the second-most, and so on.

Consider the situation faced by a given candidate i when he chooses his offer distribution, assuming that each other candidate will use the equilibrium offer distribution. Let $F(x)$ denote the cumulative probability that any given voter will be offered less than x by any other given candidate, according to this equilibrium distribution. When candidate i offers x to a voter, the probability that this candidate i will be ranked in position j by this voter is $P(j, F(x))$, where we let

$$P(j, q) = q^{K-j} (1 - q)^{j-1} (K - 1)! / ((K - j)! (j - 1)!).$$

(Here "!" denotes the factorial operator, so $(K - 1)! = 1 \times 2 \times \dots \times (K - 1)$.) That is, $P(j, q)$ denotes the probability that exactly $K - j$ of the $K - 1$ other candidates will offer the voter less than x , given that each other candidate has an independent probability q of offering less than x to this voter. Thus, if candidate i offers x to a voter, then the expected value of the points that this voter will give to this candidate is $R(F(x))$, where we let

$$R(q) = \sum_{j=1}^K P(j, q) s_j.$$

(Things could be slightly more complicated if there were a positive probability of other candidates offering exactly x , but we can ignore such complications

here. We will show in the proof of Theorem 2 that the equilibrium distribution cannot assign positive probability to any single point.)

When all candidates independently use the same offer distribution, they must all get the same expected score, which must equal \bar{s} , the average point score given by each voter. So there is a symmetric equilibrium in which all candidates use the cumulative distribution F if and only if, for any other distribution G that is on the nonnegative numbers and has mean 1,

$$\int_0^{\infty} R(F(x)) dG(x) \leq \bar{s}.$$

Theorem 2. In a K-candidate election under the rank-scoring rule with ranking points (s_1, s_2, \dots, s_K) , there is a unique symmetric equilibrium of the candidates' offer-distribution game. In this equilibrium, each candidate chooses to generate offers according to a distribution that has support on the interval from 0 to $1/\bar{s}$, and which has a cumulative distribution $F(\cdot)$ that satisfies the equation

$$x = R(F(x))/\bar{s}, \quad \forall x \in [0, 1/\bar{s}].$$

Proof. To find an equilibrium, we show first that the equilibrium offer distribution must be continuous, that is, it cannot have any points of positive probability. If all candidates used an offer distribution that assigned a positive probability δ to some point x , then there would be a positive fraction (δ^K) of the voters who would be exactly indifferent among the two candidates. Any candidate could then move his average point score among this block from \bar{s} up to s_1 by giving an arbitrarily small increase (say, ϵ) to most of the voters to whom he was going to offer x ; and the cost of this increase ($\epsilon\delta$) could be financed by moving an arbitrarily small fraction (ϵ/x) of this block down to zero. That is, if the distribution had a positive mass at some point, then a

candidate could gain a positive block of votes by a transfer of resources that would lower his score from only an arbitrarily small number of voters.

Applying the definitions of $P(\cdot, \cdot)$ and $R(\cdot)$, we now show that

$$R(0) = s_K = 0, \quad R(1) = s_1 = 1,$$

and $R(\cdot)$ is a continuous and strictly increasing function over the interval from 0 to 1. These equations hold because $P(j, 0)$ equals 0 unless j equals K , $P(j, 1)$ equals 0 unless j equals 1, and $P(K, 0) = 1 = P(1, 1)$. Continuity of $R(\cdot)$ follows directly from the formulas, because $R(q)$ is a polynomial in q . To show that $R(\cdot)$ is increasing, first verify that

$$R(q) = \sum_{j=2}^K (s_{j-1} - s_j) \sum_{m < j} P(m, q),$$

using $s_K = 0$. Then observe that

$$\sum_{m < j} P(m, q)$$

denotes the probability that more than $K - j$ other candidates have made offers in an interval of probability q , and this probability must be a strictly increasing function of q . The ordering of the s_j values guarantees that at least one term in this $R(q)$ expression must have a positive $(s_{j-1} - s_j)$ coefficient, and none can be negative.

Next, we show that the lowest permissible offer 0 must be in the support of the equilibrium distribution of offers. The essential idea is that, if the minimum of the support were strictly greater than zero, then a candidate would be devoting positive resources to voters near to the minimum of the support of the distribution. He would expect to get almost no votes ($s_K = 0$) from these voters, because all other candidates would almost surely be promising them more. Thus, it would be better to reduce the offers to 0 for most of these voters in order to make a serious bid for at least some of their votes.

The above argument can be formalized as follows. Because there are no

points of positive probability, the cumulative distribution $F(\cdot)$ is continuous. Let z denote the minimum of the support of the equilibrium distribution, so $F(z) = 0$ but $F(z + \epsilon) > 0$ for all positive ϵ . Now, select any fixed y such that $y > z$ and $F(y) > 0$. For any ϵ such that $0 < \epsilon < y - z$, a candidate might consider deviating from equilibrium by promising either y or 0 to each voter in the block of voters whom he was supposed to offer between z and $z + \epsilon$, according to his F -distributed random-offer generator. These voters in this block were going to be given offers that averaged some amount between z and $z + \epsilon$, so he can offer y dollars to at least a z/y fraction of these voters without changing his offers to any other voters. Among this z/y fraction of the block, he would get an average point score of $R(F(y))$, by outbidding the other candidates who are using the F distribution; so the deviation would get him an average point score of at least $(z/y)R(F(y))$ from this block of voters. (The voters moved down to zero in this deviation would give him $s_k = 0$ points.) If he follows the equilibrium, however, he gets at most $R(F(z + \epsilon))$ as his average point score from this block of voters. So to deter such a deviation, we must have $(z/y)R(F(y)) \leq R(F(z + \epsilon))$, and so $z \leq yR(F(z + \epsilon))/R(F(y))$. But $R(F(z + \epsilon))$ goes to $R(F(z)) = R(0) = 0$ as ϵ goes to 0 , and so z must equal 0 .

Now, let x and y be any two numbers in the support of the equilibrium distribution such that $0 < x < y$. A candidate could deviate by taking a block of voters to whom he is supposed to give offers close to x , according to his equilibrium plan, and instead he could give them offers close to y to an x/y fraction of this block and he could offer 0 to the remaining $(1 - x/y)$ fraction. Because the support of the distribution contains 0 as well as x and y , neither this self-financing deviation nor its reverse (offering close to

x to a block of voters of whom an x/y fraction were supposed get close to y , and the remaining $(1 - x/y)$ fraction were supposed to get close to 0) should increase the candidate's expected average point score from this block of voters. Thus, we must have

$$R(F(x)) = (x/y)R(F(y)) + (1 - x/y)R(F(0)).$$

But $R(F(0)) = R(0) = 0$, so $R(F(x))/x = R(F(y))/y$ for all x and y in the support of the equilibrium distribution. So there is some positive constant α such that, for all x in the support of the distribution,

$$R(F(x)) = \alpha x.$$

The maximum of the support, where $F(x) = 1$, must be where $\alpha x = R(1) = 1$; so the maximum of the support must be at $1/\alpha$.

The support is the region where the cumulative distribution function is increasing, and F has no discontinuous jumps, so $F(x)$ can equal the strictly increasing function $R^{-1}(\alpha x)$ on the entire support only if the support is the connected interval from 0 to the upper bound $1/\alpha$, with no breaks in between. (The inverse function $R^{-1}(\cdot)$ is well-defined and strictly increasing on $[0,1]$ because $R(\cdot)$ is a strictly increasing function from $[0,1]$ onto $[0,1]$.)

To evaluate the constant α , we use the fact that the mean offer must equal 1 under the F distribution, so

$$\int_0^{1/\alpha} x \, dF(x) = 1.$$

We also know that a candidate who uses the same offer distribution F as all the other candidates must expect the average point score \bar{s} , so

$$\bar{s} = \int_0^{1/\alpha} R(F(x)) \, dF(x) = \int_0^{1/\alpha} \alpha x \, dF(x) = \alpha.$$

Thus, the support of the F distribution is the interval from 0 to $1/\alpha = 1/\bar{s}$, and the cumulative distribution satisfies the formula

$$R(F(x)) = \bar{s}x, \quad \forall x \in [0, 1/\bar{s}].$$

It only remains to verify that we do get an equilibrium when F satisfies this formula. In general, for any nonnegative x , we have $R(F(x)) \leq \bar{s}x$, because $R(F(x)) = R(1) = 1 < \bar{s}x$ when $x > 1/\bar{s}$. So using any other distribution G that has mean 1 and is on the nonnegative numbers, would give a candidate an expected score

$$\int_0^{\infty} R(F(x)) dG(x) \leq \int_0^{\infty} \bar{s}x dG(x) = \bar{s},$$

with equality if the support of G is contained in the interval $[0, 1/\bar{s}]$. Thus, no candidate can increase his expected score by deviating from F to some other distribution, when all other candidates are using the distribution F .

Q.E.D.

As a simple measure of potential inequality, the maximal offer may be defined as the highest offer in the support of a symmetric equilibrium offer distribution, for any given electoral system with any given number of candidates. By Theorem 2, this maximal offer is just the reciprocal of \bar{s} , the average of the ranking points, for any rank-scoring rule. Given any K , this average \bar{s} is smallest under plurality voting, in which each voter gets a single nontransferable vote, because we then have $s_j = 0$ for all $j > 1$.

Thus, for any given number of candidates, plurality voting generates the most inequality of offers in a symmetric equilibrium, in the sense of maximizing the maximal offer that candidates give any voters. (Of course, we should not necessarily assume that the number of candidates for an office would remain constant when the electoral system is changed.) For plurality voting,

$$R(q) = P(1, q) = q^{K-1} \quad \text{and} \quad \bar{s} = 1/K,$$

So the equilibrium cumulative distribution satisfies

$$x = (F(x))^{K-1} / (1/K) = K(F(x))^{K-1}, \quad \forall x \in [0, 1/(1/K)] = [0, K],$$

and so,

$$F(x) = (x/K)^{1/(K-1)}, \quad \forall x \in [0, K].$$

When there are 4 candidates under plurality voting, each candidate offers less than 0.50 to half of the voters, and offers less than 1 to 65% of the voters, but makes offers more than 2 to 20% of the voters, with some voters getting offers as large as 4. When the number of candidates increases to 10, then each candidate offers less than 0.02 to half of the voters, and favors a 16% minority with high offers above 2, some of whom get as much as the maximal offer 10. In general, the symmetric equilibrium under plurality voting has a maximal offer of K which is larger than the median offer of $K(.5^{K-1})$ by a factor of 2^{K-1} .

Each candidate i knows that, when the other $K - 1$ candidates use the equilibrium distribution for plurality rule, the probability of his being ranked first by a voter to whom he offers x is

$$F(x)^{K-1} = ((x/K)^{1/(K-1)})^{K-1} = x/K.$$

That is, the highest of the other $K - 1$ candidates' offers, which candidate i must exceed if he is to win the voter's vote, has a uniform distribution over the interval from 0 to K in equilibrium. Because of this uniform distribution, adding an extra ε units of money to the offer of any one voter has as much chance of gaining his vote, no matter what this original offer might have been (between 0 and $K - \varepsilon$). Thus, there is no incentive for the candidate to revise his offers by taking from one voter and giving to another.

Among rank-scoring rules, for any given number of candidates K , the lowest possible maximal offer is $K/(K - 1)$, which is achieved by negative-plurality voting, where $s_j = 1$ for all $j < K$. In this sense, negative-plurality voting gives us the most egalitarian offer distributions. In equilibrium under negative-plurality voting, no voter is offered more than 1.34 when there are 4

candidates, and no voter is offered more than 1.12 when there are 10 candidates.

Following the insightful analysis of Cox (1987), we may characterize negative-plurality voting as a last-place punishing scheme, whereas (ordinary) plurality voting is a first-place rewarding scheme. Under negative-plurality voting, each candidate wants to minimize the fraction of the voters who think that he is the worst candidate, so candidates try to avoid offending many voters. Each candidate makes low offers close to 0 to a small minority of voters, in order to finance competitive offers above 1 to the majority, but the number of such neglected voters is kept small. For negative-plurality voting, the equation in Theorem 2 becomes

$$x = (1 - (1 - F(x))^{K-1}) / ((K - 1)/K)$$

This formula implies, for example, that each candidate offers less than 1 to only 37% of the voters when there are 4 candidates, and each candidate offers less than 1 to only 23% of the voters when there are 10 candidates.

Under a voting rule in which each voter must vote for exactly V of the K candidates, the maximal offer is K/V , because $\bar{s} = V/K$. Using the general formula in Theorem 2 to compute the inverse of the cumulative distribution function, standard deviations of the equilibrium offer distributions can also be calculated. Table 1 shows such these standard deviations for voting systems with V noncumulative votes and K candidates. Notice that, like the maximal offer K/V , these standard deviations increase as K increases and as V decreases. In each row of Table 1, the leftmost entry is for plurality voting and the rightmost entry is for negative-plurality voting. The probability densities of the equilibrium offer distributions, when there are 4 candidates, are shown for the cases of $V = 1$ (plurality), $V = 2$, and $V = 3$

(negative-plurality) in Figures 1, 2, and 3 respectively.

[INSERT TABLE 1 AND FIGURES 1, 2, 3 ABOUT HERE]

Borda voting is another important rank-scoring rule, in which

$$s_j = (K - j)/(K - 1).$$

So $\bar{s} = 1/2$ and the maximal offer is 2. The function $R(F(x))$ under Borda rule is the expected value of $(K - j)/(K - 1)$, where j is a random variable such that $K - j$ has a binomial distribution with expected value $(K - 1)F(x)$. Thus, by Theorem 2,

$$\begin{aligned} x/2 = \bar{s}x = R(F(x)) &= \sum_{j=1}^K P(j, F(x))(K - j)/(K - 1) \\ &= (K - 1)F(x)/(K - 1) = F(x), \quad \forall x \in [0, 2]. \end{aligned}$$

So for any number of candidates K , the equilibrium offer distribution under Borda rule is a uniform distribution over the interval from 0 to 2. The standard deviation of this distribution is approximately .58.

For a voting rule which has the same maximal offer as Borda but which has a tighter distribution as measured by the standard deviation, consider the rank-scoring rule in which $s_1 = 1$, $s_j = 1/2$ for all j such that $1 < j < K$, and $s_K = 0$. (This rule is equivalent to giving each voter one positive vote and one negative vote.) Figure 4 shows the offer distribution for this rule when $K = 4$. The maximal offer under this voting rule is also equal to 2 for all K , because $\bar{s} = 1/2$, but the equilibrium offer distribution has standard deviation .52 when $K = 4$, and .37 when $K = 8$.

[INSERT FIGURE 4 ABOUT HERE]

As we compare these equilibrium offer distributions, the fact that plurality voting generates much greater inequality among voters than other

electoral systems may be quite disturbing. In defence of plurality voting, one may argue that our symmetry assumption is unlikely to be satisfied in plurality elections with many candidates. Duverger's law (see Riker 1982) predicts that, in multicandidate races under plurality voting, the voters will take only two candidates seriously, and each voter will vote for the candidate whom he prefers among these two serious candidates (even if there are other candidates whom he likes even better). That is, the symmetric equilibrium in which every candidate is considered to have a serious chance of winning is much less likely to be observed than one of the nonsymmetric equilibria in which only two candidates are considered to have a substantial chance of winning. By showing the problematic nature of the symmetric equilibrium that Duverger's law excludes, the analysis here casts a rather favorable light on Duverger's law. If Duverger's law did not hold, then the use of plurality voting in multicandidate winner-take-all elections would more often give us elected leaders who deliberately appeal to only a small minority of the voters.

Approval voting

Under approval voting, each voter indicates approval of a set of candidates on his ballot. A voter is allowed to approve of any number of candidates. Each candidate's approval score is then the fraction of the voters who have approved of him. That is, each voter can give each candidate either an approval worth 1 point, or a disapproval worth 0 points, and the winner is the candidate who gets highest average point score from the voters. Approval voting is a scoring rule but it is not a rank-scoring rule, because the point scores of various candidates cannot be determined simply from a voter's rank-ordering of the candidates. For example, when there are three candidates,

one voter might give an approval vote only to his one favorite candidate, whereas another voter with the same rank-ordering of candidates might give an approval vote to each of his top two candidates.

We need to formulate a voter's decision criterion, to describe how each voter would decide how many candidates to approve, in a symmetric equilibrium under approval voting. Let us number the K candidates $1, 2, \dots, K$. Now consider a voter, and let u_i denote the value of the offer that candidate i has promised to this voter. The voter's decision to give an approval vote to candidate i will only matter if candidate i is in a close race with some other candidate h (where h would win unless this voter adds an approval vote for i). In such a case, giving an approval vote to candidate i would change the voter's payoff from u_h to u_i . Our symmetry assumption implies that, in the case of such a close race involving candidate i , any of the other $K - 1$ candidates is equally likely to be the candidate h who is in a close race with candidate i . Thus, the expected net gain for the voter from giving an approval vote to candidate i is proportional to

$$\sum_{h \neq i} (u_i - u_h) / (K - 1) = (u_i - \bar{u})(K / (K - 1)),$$

where

$$\bar{u} = \sum_{h=1}^K u_h / K.$$

Thus, in a symmetric equilibrium, each voter should compute the average \bar{u} of the offers that the K candidates have promised him, and he should give an approval vote to every candidate i such that $u_i > \bar{u}$, and he should not approve of any candidate who has offered him less than this average \bar{u} .

Given any m candidates who are generating their offers independently according to the cumulative distribution function $F(\cdot)$, let $A_m(x)$ denote the probability that the average of their m offers is less than x , for any given

voter.⁴ Now, suppose that some candidate i offers an amount u_i to some voter, while the other $K - 1$ candidates are selecting their offers independently according to the cumulative distribution function F . Then the probability of candidate i getting an approval vote from this voter is $A_{K-1}(u_i)$, because candidate i 's offer u_i is greater than the average of all K offers if and only if u_i is greater than the average of the other $K - 1$ candidates' offers.

To win, a candidate wants both to increase his own score and decrease the scores of his opponents. Under approval voting, increasing the vote score for candidate i is not the same thing as decreasing the vote scores of the other candidates, because the total number of points that each voter gives is not fixed under approval voting (unlike the rank-scoring rules that we considered in the preceding section). For any candidate h other than i , let $B(u_i)$ denote the probability that candidate h will get an approval vote from any given voter, when candidate i offers u_i and all other candidates' offers are drawn independently from the F distribution.⁵

Candidate i would gain by deviating from the F distribution to some other distribution G , when all other candidates are expected to use the F distribution, if the deviation would cause candidate i 's average approval score to become higher than the other candidates' average approval scores. We are assuming here that the number of voters is large or effectively infinite, so that a candidate's average approval score is equal to the probability of any one voter giving him an approval vote. Thus, there is an equilibrium in which all candidates generate offers according to the F distribution iff

$$\int_0^{\infty} A_{K-1}(x) dG(x) \leq \int_0^{\infty} B(x) dG(x)$$

for every nonnegative distribution G with mean 1. (Unwinding the definitions, it can be shown that this inequality becomes an equality when $G = F$.)

It seems quite difficult to explicitly calculate equilibrium offer distributions for approval voting.⁶ I have analyzed approximate solutions for a discrete version of this problem, in which all offers are integer multiples of 0.05. The resulting equilibrium probability distribution for the case of 4 candidates is shown in Figure 5. For any number of candidates up to 10, Table 2 shows the standard deviation of the equilibrium offer distribution, the maximal offer, and the average approval score for each candidate, in the symmetric equilibria that have been found for this .05-discrete approximation. In each case, I have found only one equilibrium for each number of candidates. The convergence of the algorithm seemed sufficiently robust to justify a conjecture that this symmetric equilibrium may be unique, but I cannot prove that there is a unique symmetric equilibrium under approval voting.

{INSERT FIGURE 5 AND TABLE 2 ABOUT HERE}

As Table 2 shows, equilibrium offer distributions for approval voting have maximal offers that are between 1 and 2, and these maximal offers decrease as the number of candidates increases. As Figure 5 illustrates, the support of the equilibrium offer distribution is not convex when $K \geq 3$. Instead, each candidate gives offers that are greater than 1 to a majority of the voters, while the remaining voters (a minority which decreases in size as K increases) get offers that are close to 0.

To get some intuition for these results, suppose that the number of candidates K is large, and consider a nonequilibrium scenario in which all candidates draw their offers from the uniform distribution on 0 to 2. The central limit theorem tells us that the average of any $K - 1$ candidates' offers to any given voter has a distribution that is approximately normal, with mean 1

and a small standard deviation. The probability that a candidate would gain (or lose) a voter's approval by slightly increasing (or decreasing) this voter's offer is proportional to the probability density of the average of the other $K - 1$ candidates' offers, and the normal approximation has a density that is highest at 1. So a candidate could increase his expected approval score by decreasing offers that are far from 1, and then using the resources thus made available to correspondingly increase offers that are close to 1. The result of such transfers would be to move towards a distribution in which some voters get offers near 0, and the rest of the voters get offers are larger than 1 but less than 2. Such a distribution is in fact shown in Figure 5, where about 31% of the voters get offers close to 0, 25% of the voters get offers close to 1.15, 15% of the voters get offers close to 1.50, and 29% of the voters get offers close to 1.70. It can be shown that, like the rank-scoring rules of the preceding section, the equilibrium offer distribution under approval voting must be continuous, even though the support of the distribution is not convex.

Single transferable vote

Let us consider now a simple version of single transferable vote (STV), in which K candidates are competing for a single office. (Multiseat councils will be considered in the next section.) Each voter submits a ballot that ranks the K candidates in some order, chosen by the voter. These ballots are then recounted a sequence of times until a candidate wins. In the first count, each voter's ballot is counted as a vote for the candidate who is listed at the top of the voter's rank ordering. If no one candidate gets a majority in this count, then the candidate with the lowest share of the votes is eliminated from consideration, and each ballot that was assigned to this candidate is

reassigned to the candidate who is ranked highest on the ballot among those who have not yet been eliminated. Then the ballots are recounted, and this process is repeated until some candidate gets a majority. When a candidate gets a majority of the votes in a count, then that candidate wins the election. Thus, STV is a ranking rule, in the sense that each voter's ballot can be written as a rank-ordering of the candidates, but it is not a scoring rule, because a voter's rank-ordering does not simply generate one point score for each candidate.

Our symmetry assumption implies that voters will vote sincerely under STV (because a voter will not perceive any pair of candidates as being more likely than any other pair to be in a close race where his vote could make a difference in the outcome). That is, we may assume here that each voter will rank the candidates in order of their offers to him, putting the candidate who has offered the most at the top of his rank ordering.

In a symmetric equilibrium of our model, we should expect the candidates to divide the votes equally at every round, and so we need to be careful about the interpretation of ties. Remember that our model assumes that the number of voters is infinite, but this assumption is meant only as an approximation to a large finite population. When two or more candidates get the same share of the votes in our model, we should understand that this means that the difference in their vote totals is a vanishingly small or infinitesimal fraction of the overall population, but (because the population is very large) there may still be hundreds of votes difference between the absolute vote totals of these candidates. Thus, a tie for last place in our infinite-population model should be interpreted as an outcome in which exactly one of these candidates is actually the lowest. By symmetry of the scenario, each of the apparently tied

candidates is equally likely to be the actual last-place candidate who then is eliminated.

When all candidates independently generate offers from the same cumulative distribution F , there will be a symmetric allocation of votes, and so no candidate will ever get a majority until there have been $K - 1$ counts, at the last of which all but two candidates will have been eliminated (and so each will get an approximate 50% share, which will really be a strict majority for one of the two candidates). At the first count, each candidate will have a probability $1/K$ of being eliminated, and a probability $(K - 1)/K$ of being retained for the second count. At the second count, each remaining candidate will have a probability $1/(K - 1)$ of being eliminated, and a probability $(K - 2)/(K - 1)$ of being retained; and so on. Each candidate's probability of winning is thus

$$((K - 1)/K)((K - 2)/(K - 1)) \dots (1/2) = 1/K.$$

Now suppose that some candidate i deviates from this symmetric scenario and generates offers according to some other cumulative distribution G , while all other candidates continue to use the F distribution. Consider a count in which candidate i is one of n candidates who have not yet been eliminated. A voter who has been offered u_i by candidate i will rank candidate i at the top of his list, when the eliminated candidates are ignored, if candidate i offered him more than the other $n - 1$ candidates, which has probability $F(u_i)^{n-1}$. So candidate i 's vote share at this count will be

$$\int_0^{\infty} (F(u_i))^{n-1} dG(u_i).$$

If this vote share is greater than the average $1/n$, then candidate i will surely not be eliminated at this count. On the other hand, if this vote share at this round is less than $1/n$, then candidate i must be in last place, because

the other candidates are symmetric in this scenario, and so candidate i will surely be eliminated at this count.⁷

Being strictly higher than everyone else in the early counts is irrelevant if elimination is certain at any one of the later counts. So a deviating candidate's probability of winning will go to zero if his expected vote share is less than the average at any of the $K - 1$ counts of the ballots. Thus, we may say that there is an equilibrium in which all candidates use the F distribution iff there does not exist any probability distribution G on the nonnegative real numbers, with mean 1, such that

$$\int_0^{\infty} (F(x))^{n-1} dG(x) \geq 1/n,$$

for every n in $\{2, 3, \dots, K\}$, and

$$\int_0^{\infty} (F(x))^{m-1} dG(x) > 1/m,$$

for at least one number m in $\{2, 3, \dots, K\}$.

We can now show that there are multiple symmetric equilibria under STV. These equilibria include offer distributions that are arbitrarily close to any of the equilibria that we got for plurality voting when the number of candidates was between 2 and K . Thus, any number between 2 and K can be the maximum of a symmetric equilibrium offer distribution under STV. The intuition behind this multiple-equilibrium result is simple. To gain from deviating from a symmetric scenario, a candidate must avoid losing in $K - 1$ counts, and so an equilibrium can be supported by any profile of campaign strategies that deters such deviations in at least one of these $K - 1$ counts.

Theorem 3. Let $(\lambda_2, \dots, \lambda_{K-1}, \lambda_K)$ be any vector such that

$$\lambda_n > 0, \quad \forall n \in \{2, \dots, K - 1, K\}, \quad \text{and} \quad \sum_{m=2}^K \lambda_m = 1.$$

Let F be defined by the equation

$$x = \sum_{n=2}^K n \lambda_n (F(x))^{n-1}, \quad \forall x \in [0, \sum_{m=2}^K m \lambda_m].$$

Then there is an equilibrium under single transferable vote in which all candidates generate offers according to the F distribution.

Proof. Notice first that the function F described in the theorem always exists, because the equation directly defines its inverse

$$F^{-1}(q) = \sum_{n=2}^K n\lambda_n q^{n-1}, \quad \forall q \in [0, 1].$$

Then F is well defined because F^{-1} is strictly increasing and thus one-to-one on its range. To verify that the mean is 1, observe that

$$\begin{aligned} \int_0^{\infty} x \, dF(x) &= \int_0^1 F^{-1}(q) \, dq = \int_0^1 \left(\sum_{n=2}^K n\lambda_n q^{n-1} \right) \, dq \\ &= \sum_{n=2}^K n\lambda_n \left(\int_0^1 q^{n-1} \, dq \right) = \sum_{n=2}^K n\lambda_n (1/n) = 1. \end{aligned}$$

If the F distribution did not form an equilibrium, then there would exist an offer distribution G such that

$$\sum_{n=2}^K \lambda_n \left(\int_0^{\infty} (F(x))^{n-1} \, dG(x) \right) > 1$$

because the coefficient of λ_n would be at least 1 in each term, with at least one strict inequality. Then, by the definition of F, we would get

$$1 < \int_0^{\infty} \left(\sum_{n=2}^K n\lambda_n (F(x))^{n-1} \right) \, dG(x) = \int_0^{\infty} x \, dG(x),$$

and so the deviation G must have a mean greater than 1, which is not permissible for a candidate. Q.E.D.

In the extreme case where some λ_n coefficient is close to 1 and all other λ_m are close to 0, the equation in Theorem 3 that defines F becomes close to $x = n(F(x))^{n-1}$, which characterizes the symmetric equilibrium offer distribution for plurality voting with n candidates.

Multiseat elections

We have until now only discussed the case of individual candidates competing for a single office, but most of our results can be readily extended

to competition for multiple seats and to competition among parties. The crucial defining property of the symmetric equilibria that we have studied is that no candidate can break the symmetry in his own favor (scoring higher than the still-symmetric other candidates) by deviating from the equilibrium strategy. This characterization of symmetric equilibria remains the same when there are many seats at stake, as long as the number of candidates is greater than the number of seats.

For example, consider an election in which K candidates are competing for L seats in a council, where $1 \leq L \leq K - 1$, and a rank-scoring rule is used with ranking points (s_1, s_2, \dots, s_K) such that $1 = s_1 \geq s_2 \geq \dots \geq s_K = 0$. The candidates with the L highest scores will win the L seats. Suppose, for simplicity, that each candidate who wins a seat in the council will control a portion of the government budget that has an expected average value of one dollar per voter. Each winner will allocate his portion of the government budget to the voters according to the campaign promises that he independently generated from his offer distribution. (A more complicated story might say that a member of the council will get a share of the budget only if he is a member of a governing majority coalition that will take control of the council after the election. When the candidates are symmetric, however, we may assume that any candidate who wins a seat will be equally likely to be included in the governing coalition, and so each candidate has the same expected budget to allocate if he wins.)

Under these assumptions, the symmetric equilibrium offer distribution, for any given rank-scoring rule, is exactly the same as described in Theorem 2. In this multiseat case, each candidate has a probability L/K of winning in the symmetric equilibrium, but (as in the 1-seat case) he still would deviate only

if doing so would raise his expected score above the other still-equal (by symmetry) scores of the other $K - 1$ candidates. So the conditions for a symmetric equilibrium with K candidates depend only on the ranking points (s_1, s_2, \dots, s_K) but do not depend on the number of seats L , as long as $L < K$. Thus, our results for the 1-seat case in Theorem 2 can be immediately generalized to multiple seats. For example, the rank-scoring rule with the ranking-points vector $(1, 0, \dots, 0)$ is called single nontransferable vote or SNTV when there are multiple seats at stake. So the symmetric equilibrium offer distribution for SNTV with K candidates and L seats is the same as the symmetric equilibrium offer distribution that we found for plurality voting with K candidates in the 1-seat case.

Similarly, the equilibrium offer distributions that we described for approval voting with K candidates and 1 seat, apply equally well when the K candidates are competing for L seats and the winners will be the candidates with the L highest approval scores.

Under single transferable vote, we do find some difference in the case of multiple seats. With L seats, the majority quota is replaced by a $1/(L + 1)$ quota, and the recounting stops after the count in which one candidate is eliminated out of $L + 1$ remaining. Thus, we need to revise Theorem 3 by changing the number "2" to " $L + 1$ " throughout, but then the revised theorem can be applied to the case where L seats are being allocated under STV. So when K candidates compete for L seats under STV, there are multiple symmetric equilibria in which the maximal offers range from $L + 1$ up to K .

The most substantial difference that we may find in the multiseat case is that the assumption of a symmetric multicandidate equilibrium may become more compelling. In single-seat plurality elections, Duverger's law predicts that

two candidates will be distinguished as the only two serious contenders. In a multiseat election under SNTV, however, we should expect that there will be more serious candidates than seats (see Reed 1990 and Cox 1992). Thus, with multiple seats, we cannot use Duverger's law to escape from the fact that single nontransferable vote may encourage candidates to make extreme promises to narrow subsets of the electorate.

SNTV is used in multiseat districts in Japan, and it is also used in intraparty competition in some open-list proportional representations systems, as in Brazil. In open-list PR systems, when a party nominates more candidates than the number of seats that it wins, the nominated candidates who actually get the party's seats are determined by voters' votes for the individual candidates. To be specific, suppose that a party has nominated K candidates, but the party is expected to win only L seats, and each voter who votes for the party can vote for one individual candidate on his ballot. The party's L seats will be won by those among its candidates who get the L highest scores, in this contest among the individual candidates. Then our analysis of plurality voting and SNTV with K candidates can be directly applied to the race among the K nominees of this party for its L seats. Our results suggest that the K nominees will make highly unequal offers to voters, and each will concentrate his attention on a small minority of the voters who support the party. This prediction is consistent with the bailiwick phenomena (in which candidates seek votes only in very narrow subsets of the electoral district) that Ames (1992) has described in Brazilian legislative elections. Our model predicts that the candidates could be induced to substantially broaden their appeal if the individual candidates in party lists were ranked by another voting system, such as Borda or multiple noncumulative votes or approval voting.

The analysis of this paper can also be applied to the competition among parties in proportional representation systems. For simplicity, suppose that each party may expect to control a portion of the government budget that is proportional to the share of the legislature that it wins. Suppose also that each party will allocate its portion of the budget to individual voters in proportion to its campaign promises, which it has independently generated for every voter according to its offer distribution. Each voter must choose one party to support, and each party gets seats in proportion to its share of the vote. So each voter should support the party that has promised him the most.

Under these assumptions, our analysis of proportional representation looks like that of plurality voting and SNTV, because proportional representation uses the same ranking-points vector $(1,0,\dots,0)$ and thus is also a pure first-place rewarding scheme. The symmetric equilibrium offer distributions in a proportional representation election among K parties will be the same as the symmetric equilibrium offer distributions in a 1-seat plurality election among K candidates. Thus, in a proportional representation system, if there are many parties that are perceived as having similar potential to participate in the process of forming a government, then each of these parties may feel compelled to narrowly appeal to a small group of voters, and their campaign promises may tend to create special-interest divisions even in an otherwise homogeneous electorate.

Minority representation

This paper has focused on only one attribute of electoral systems: the degree to which they may encourage candidates to create inequalities among the voters by favoring some voters at the expense of others. For a full comparison

of electoral systems, we must consider other models that emphasize other attributes of electoral systems. One other attribute that we should at least briefly consider here, because of its close relationship to the main topic of this paper, is the ability of electoral systems to guarantee diverse representation of minority groups. We follow here the seminal analysis of this question by Cox (1987, 1990).

Consider again an election in which K candidates are competing for election to a council of L members, and people vote according to a rank-scoring rule with ranking points (s_1, s_2, \dots, s_K) such that $1 = s_1 \geq s_2 \geq \dots \geq s_K = 0$. We assume that $L \leq K - 1$, and the candidates with the L highest scores will win the L seats. Suppose now that (instead of making a discretionary allocation of a given budget, as has been assumed in this paper up until now) government officials have only one simple Yes-or-No policy question to decide. (E.g: "Should our nation ratify the new regional trade agreement?") At the beginning of the campaign, each candidate must choose to be either on the affirmative side or on the negative side of this policy question. Every voter has preferences one way or the other on this question, and will select his vote to try to maximize the number of elected officials who have chosen the policy answer that he prefers. As usual, we suppose that the number of voters is very large (or is an infinite measure space). Let Q denote the fraction of the voters who prefer the affirmative answer to the policy question, and suppose that Q is common knowledge.

Let us now ask, how large must the affirmative fraction Q be to guarantee that there does not exist any symmetric equilibrium in which all of candidates are expected to choose the negative side of the policy question? That is, what is the largest minority group that may be ignored by all candidates in the

election, when each candidate chooses his policy position so as to maximize his probability of winning a seat? This upper bound for the size of a voting block that all candidates could ignore may be called the Cox threshold.

If all candidates choose the negative side of the policy position, then (by symmetry) they must all expect the same average score

$$\bar{s} = \sum_{j=1}^K s_j / K,$$

and so each candidate has a probability L/K of winning a seat. If one candidate deviates to the affirmative side, while all other candidate stay on the negative side, then this deviating candidate will get a score

$$Qs_1 + (1 - Q)s_K = Q,$$

because the affirmative voters will rank him at the top of their ballots, while the negative voters will rank him at the bottom of their ballots. The symmetry assumption implies that the other $K - 1$ candidates who are all on the negative side will get the same expected score, and we know that the scores of all K candidates must average to \bar{s} . So the deviating candidate cannot expect to win if $Q < \bar{s}$, but the single deviating candidate should expect to win a seat if $Q > \bar{s}$. (The large-population assumption is used here only to justify the assumption that the candidate with the highest expected score will almost surely have the highest realized score.) So to guarantee that there is no symmetric equilibrium in which the affirmative block is ignored by all K candidates, the affirmative block must be greater than an \bar{s} fraction of the electorate. Thus, as Cox (1987, 1990) has shown, the Cox threshold for a rank-scoring rule is \bar{s} .

Of course, guaranteeing that some candidates will offer to represent the affirmative block of voters does not guarantee that the affirmative block will actually win representation in the council, except in the case of $L = K - 1$.

When there are seats for all except one candidate, then a lone candidate on one side of the policy question will fail to get a seat only if he is the lowest scorer, which cannot happen if his score Q is above the average score \bar{s} . Thus, a block of voters that is larger than \bar{s} is sure to win at least one seat in equilibrium when $L = K - 1$, even without the symmetry assumption.

For any rank-scoring rule, the Cox threshold \bar{s} is exactly the reciprocal of the maximal offer that we found in Theorem 2. In this sense, the rank-scoring rules that are best for guaranteeing representation of small minorities, in large multiseat elections, are also the worst for deterring candidates from creating inequalities in a homogeneous population. That is, electoral provisions (like SNTV) that are designed to give representation to diverse minority groups may also have the effect of encouraging candidates to create special interest groups and minority conflict even when it would not otherwise exist. Designers of electoral systems should recognize a tradeoff between these two desiderata: guaranteeing representation for minority groups in the electorate, and encouraging individual candidates to appeal broadly to the electorate as a whole.

With respect to these two desiderata, however, there seems no reason to use a system that has a low Cox threshold (which encourages a diversity of narrowly focused candidates, rather than candidates who try to appeal broadly to the whole electorate) unless the number of seats being filled is large enough for this diversity of candidates to be translated into a diversity of elected officials. In a single-seat election with two policy alternatives, for example, the minority cannot be represented by a winning candidate unless the majority is left unrepresented after the election.

To guarantee that clear majority interests are not neglected by all

candidates in equilibrium, a good electoral system should have a Cox threshold that is not greater than $1/2$. For example, negative-plurality voting with 10 candidates would give us a Cox threshold of $\bar{s} = .90$. So even if 89% of the voters preferred the affirmative position in our simple model, there would exist a symmetric equilibrium in which all 10 candidates chose the negative policy position. If one candidate deviated alone to the affirmative side, he would get a zero rating from 11% of the voters, while the nine negative candidates would each expect to get zero ratings from $.89/9 = 9.89\%$ of the voters, and so the affirmative candidate would lose. Of course, there exists another equilibrium in which all candidates coordinate on the majority side. However, the danger of having an equilibrium in which an overwhelming majority is so completely unrepresented makes negative-plurality a very undesirable system. This danger of neglecting majority interests similarly leads us to reject all rank-scoring rules that have $1/\bar{s} < 2$, even though, by Theorem 2, such rules may seem good for reducing the inequity of candidates' offer distributions.

The exact reciprocal relationship between the Cox threshold and the maximal offer which we have found for rank-scoring rules does not necessarily apply to other electoral systems. Under single transferable vote with K candidates and L seats, the Cox threshold is $1/(L + 1)$, because a candidate is guaranteed a seat when he is ranked at the top by more than $1/(L + 1)$ of the electorate. But Theorem 3 shows symmetric equilibria in which some voters get offers that are arbitrarily close to K , which can be strictly greater than $L + 1$. Thus, we find maximal offers which are above the reciprocal of the Cox threshold for STV.

Under approval voting, with any number of candidates and any number of

seats, the Cox threshold is $1/2$. If the candidates differed only in their positions on one binary policy question, then each voter would sincerely approve all candidates who adopted his preferred position; and so each candidate would maximize his approval score by choosing the position that is favored by the majority. (See Cox 1987, and Myerson 1993.) But Table 2 shows that, for approval voting with more than two candidates, the maximal offers in symmetric equilibrium offer distributions are strictly less than 2. So we find maximal offers which are below the reciprocal of the Cox threshold for approval voting, and approval voting allows us to get maximal offers less than 2 without having a Cox threshold above $1/2$.

NOTES

This research was initially inspired by conversations with Barry Ames and John Londregan. In writing this paper, I have also benefitted from many discussions with Robert Weber. Support from the Dispute Resolution Research Center at Northwestern University and from the National Science Foundation is gratefully acknowledged.

1. With a finite population of M voters, and with a fixed supply of M dollar's worth of public resources to be allocated, campaign promises could not be independent across all voters, because the offers to all voters would have to sum to the given number M . However, as the number M increases, independently distributed offers can be more and more closely approximated. If the mean of the candidate's offer distribution is not more than 1 and the support of the distribution is bounded then, for any small positive number ϵ , M voters' offers that are drawn independently from the candidate's distribution would have probability less than ϵ of totalling more than $\$M(1 + \epsilon)$, when M is sufficiently large. Thus, taking the limit as the population goes to infinity, we can assume that each candidate makes independent offers to every voter. Conversely, de Finetti's theorem in probability theory implies that, if an infinite number of voters are treated identically by a candidate, in the sense that their offers are exchangeable random variables, then their offers must be conditionally independent given the candidate's offer distribution.

2. The Riemann-Stieltjes integral is a generalization of the Riemann integral which allows us to express expected values with respect to both continuous and discrete probability distributions. Let $a(x)$ denote any bounded continuous

function of the variable x , and let $b(x)$ denote any bounded and nondecreasing function of x . Then the Riemann-Stieltjes integral

$$\int_0^{\infty} a(x) db(x)$$

can be defined as the limit, as $n \rightarrow +\infty$, of the sum

$$\sum_{i=1}^{\infty} a(i/n) \left[b(i/n) - b((i-1)/n) \right].$$

If the function b is differentiable, then

$$\int_0^{\infty} a(x) db(x) = \int_0^{\infty} a(x) b'(x) dx.$$

3. The candidates' payoff functions are discontinuous when their expected scores are equal. Simon and Zame (1990) show that, to find the equilibria of such discontinuous games, we may need to allow the payoffs to be endogenously redefined at points of discontinuity, requiring only that these endogenous payoffs should be selected in the convex-valued upper-hemicontinuous extension of the original payoff function. However, we are analyzing here only symmetric equilibria (which must exist because the game is symmetric), and candidates who use the same offer distribution must have the same probability of winning. So the methodology of Simon and Zame leads us to use a simple model in which candidates with the same expected score have the same probability of winning.

4. For each $m \geq 1$, $A_m(\cdot)$ is the cumulative probability function for the average of m offers that are sampled independently from the F distribution. These cumulative functions may be inductively calculated by the equations

$$A_1(x) = F(x),$$

$$A_m(x) = \int_0^{mx} A_{m-1}((mx - z)/(m - 1)) dF(z), \quad \forall x.$$

To derive the above formula, we use the fact that, if y is the average of $m - 1$ candidates' offers and z is another candidate's offer, then the average of

these m candidates offers is $((m - 1)y + z)/m$, which is less than x if and only if $y \leq (mx - z)/(m - 1)$.

5. To formally define $B(u_i)$, consider any other candidate h , consider some specific voter, and let y denote the average of the offers to this voter from all candidates other than i and h . Then candidate h will get an approval vote from this voter if

$$u_h > (u_h + u_i + (K - 2)y)/K.$$

So, assuming $K \geq 3$, candidate h will get an approval vote if

$$((K - 1)u_h - u_i)/(K - 2) > y.$$

Thus, the probability that candidate h will get an approval vote from this voter, when candidate i offers u_i and all other candidates' offers are drawn from the F distribution, is

$$B(u_i) = \int_0^{\infty} A_{K-2}(((K - 1)u_h - u_i)/(K - 2)) dF(u_h).$$

Here $A_{K-2}(y) = 0$ if $y < 0$.

6. In an equilibrium under approval voting, no candidate i should be able to increase the average value of $A_{K-1}(u_i) - B(u_i)$ for some block of voters by transferring promised resources among these voters. So let x and y be numbers such that $0 < x < y$ and such that x is in the support of the F distribution. Suppose that candidate i is consider the following perturbation of the F distribution: all voters who were supposed to be given an offer close to x (according to the random-number generator that implements F) will instead be offered either 0 or y , and, to make this perturbation self-financing, the fraction of these voters who will get y will be close to x/y . As a result of this perturbation, the net change in the average of candidate i 's objective

$(A_{K-1} - B)$, for this block of voters, would be

$$\begin{aligned} & (x/y)(A_{K-1}(y) - B(y)) + (1 - x/y)(A_{K-1}(0) - B(0)) - (A_{K-1}(x) - B(x)) \\ & = (C(y) - C(x))x, \end{aligned}$$

where the function $C(\cdot)$ is defined such that

$$C(z) = (A_{K-1}(z) - B(z) - A_{K-1}(0) + B(0))/z, \quad \forall z > 0.$$

This gain must be nonpositive if x is in the support of the F distribution, and it must also be nonnegative if y and 0 are in the support of the F distribution (in which case the above perturbation could be reversed). So we can construct an equilibrium offer distribution by choosing F so that the support of the F distribution is in the set of numbers z that maximize the function $C(z)$. This C -maximization condition is difficult to solve analytically, because $C(\cdot)$ depends on $F(\cdot)$, but iterative methods can be used to find approximate solutions. The Pascal program that I have used to find approximate equilibrium offer distributions for approval voting is available from the author on request.

7. There is some technical complexity in the case where the deviating candidate i 's expected vote share exactly equals the average $1/n$ in an STV recount when n candidates remain. The definition of equilibrium used here is consistent with the simple assumption that, when expected vote shares are equal, then each candidate is equally likely to be eliminated. However, if G is strictly different from F , then we cannot be sure that an equal expected vote share in the infinite-population limit implies an equal probability of being last in the case of a large finite population.

REFERENCES

- Ames, Barry. 1992. "Disparately seeking politicians: strategies and outcomes in Brazilian legislative elections." Department of Political Science, Washington University, St. Louis, MO.
- Cox, Gary W. 1987. "Electoral equilibrium under alternative voting institutions." American Journal of Political Science 31:82-108.
- Cox, Gary W. 1990. "Centripetal and centrifugal incentives in electoral systems." American Journal of Political Science 34:903-935.
- Cox, Gary W. 1992. "Strategic voting equilibria under SNTV: the M+1 rule." Department of Political Science, University of California at San Diego.
- Gross, O., and R. Wagner. 1950. "A continuous Colonel Blotto game." RAND Research Memorandum RM-408, Santa Monica, CA.
- Morton, Rebecca B., and Roger B. Myerson. 1992. "Campaign spending with impressionable voters." Northwestern University discussion paper.
- Myerson, Roger B. 1993. "Effectiveness of electoral systems for reducing government corruption: a game-theoretic analysis," Games and Economic Behavior 5:118-132.
- Myerson, Roger B., and Robert J. Weber. 1993. "A theory of voting equilibria." American Political Science Review 87:102-114.
- Owen, Guillermo. 1982. Game Theory 2nd ed. New York: Academic Press.
- Reed, Steven R. 1990. "Structure and behavior: extending Duverger's law to the Japanese case," British Journal of Political Science 20:335-356.
- Riker, William H. 1982. "The two-party system and Duverger's law," American Political Science Review 76:753-766.

Shubik, Martin. 1970. "Voting, or a price system in a competitive market structure." American Political Science Review 64:179-181.

Simon, Leo K., and William R. Zame. 1990. "Discontinuous games and endogenous sharing rules," Econometrica 58:861-872.

Number of Candidates	Number of noncumulative votes that each voter must cast						
	1	2	3	4	5	6	7
2	.58						
3	.89	.45					
4	1.13	.70	.38				
5	1.33	.89	.59	.33			
6	1.51	1.04	.75	.52	.30		
7	1.66	1.18	.89	.67	.47	.28	
8	1.81	1.31	1.01	.79	.60	.44	.26

TABLE 1. Standard deviations of the equilibrium offer distributions, when each voter must cast a given number of noncumulative votes.

Number of Candidates	Std. dev.	Max. offer	Fractn. < 1	Avg. score
2	.58	2.00	.50	.50
3	.71	1.90	.34	.55
4	.70	1.70	.31	.60
5	.66	1.60	.29	.63
6	.65	1.50	.29	.66
7	.61	1.45	.27	.69
8	.61	1.40	.27	.71
9	.56	1.35	.24	.74
10	.54	1.35	.23	.74

TABLE 2. The standard deviation of a candidate's offers, the maximal offer, the fraction of voters offered less than 1, and the average approval score for each candidate, in equilibrium under approval voting. (Calculated using a 0.05 discrete approximation.)

Figure 1. Offer distribution with 1 vote, 4 candidates.

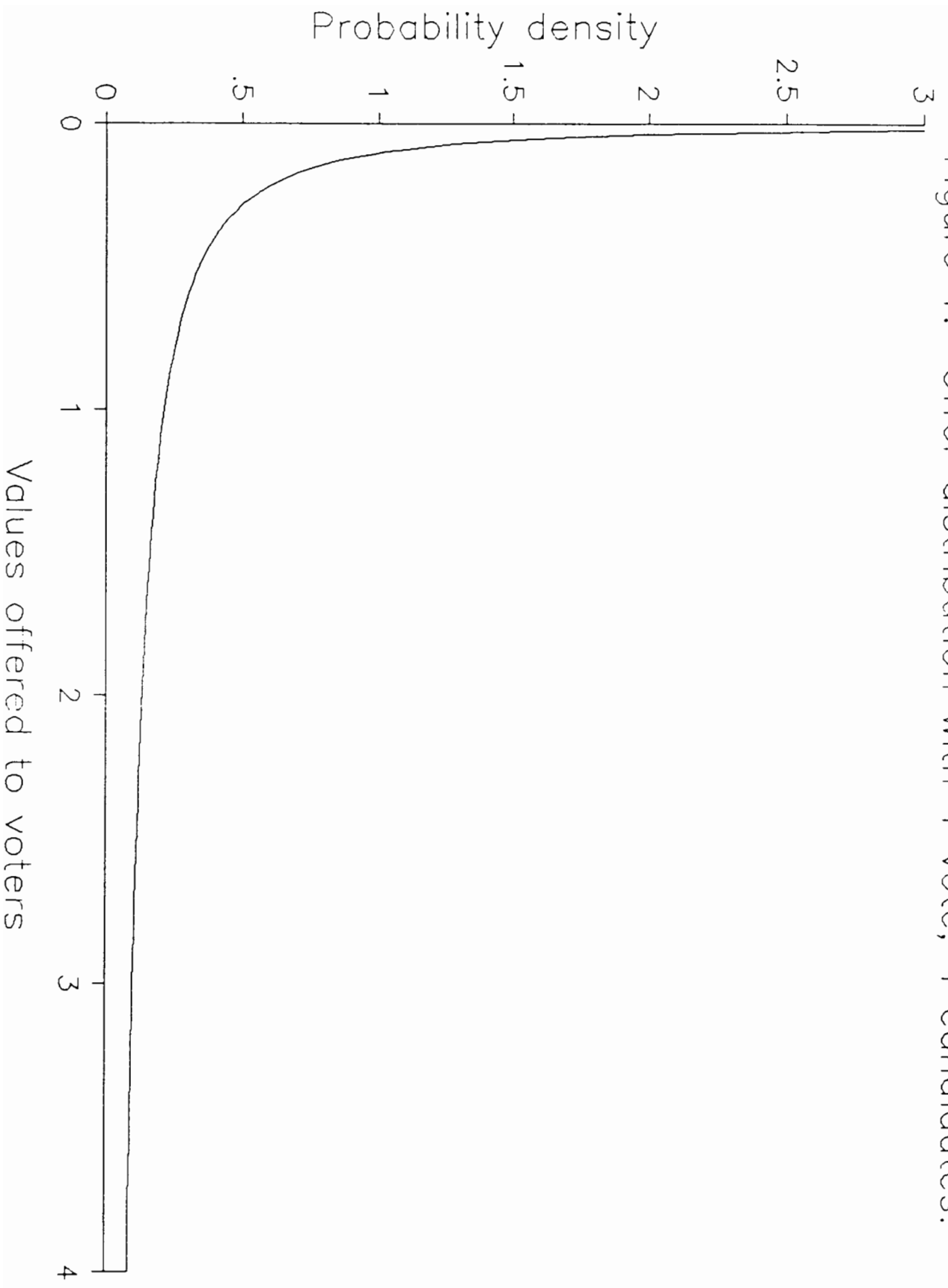


Figure 2. Offer distribution with 2 votes, 4 candidates.

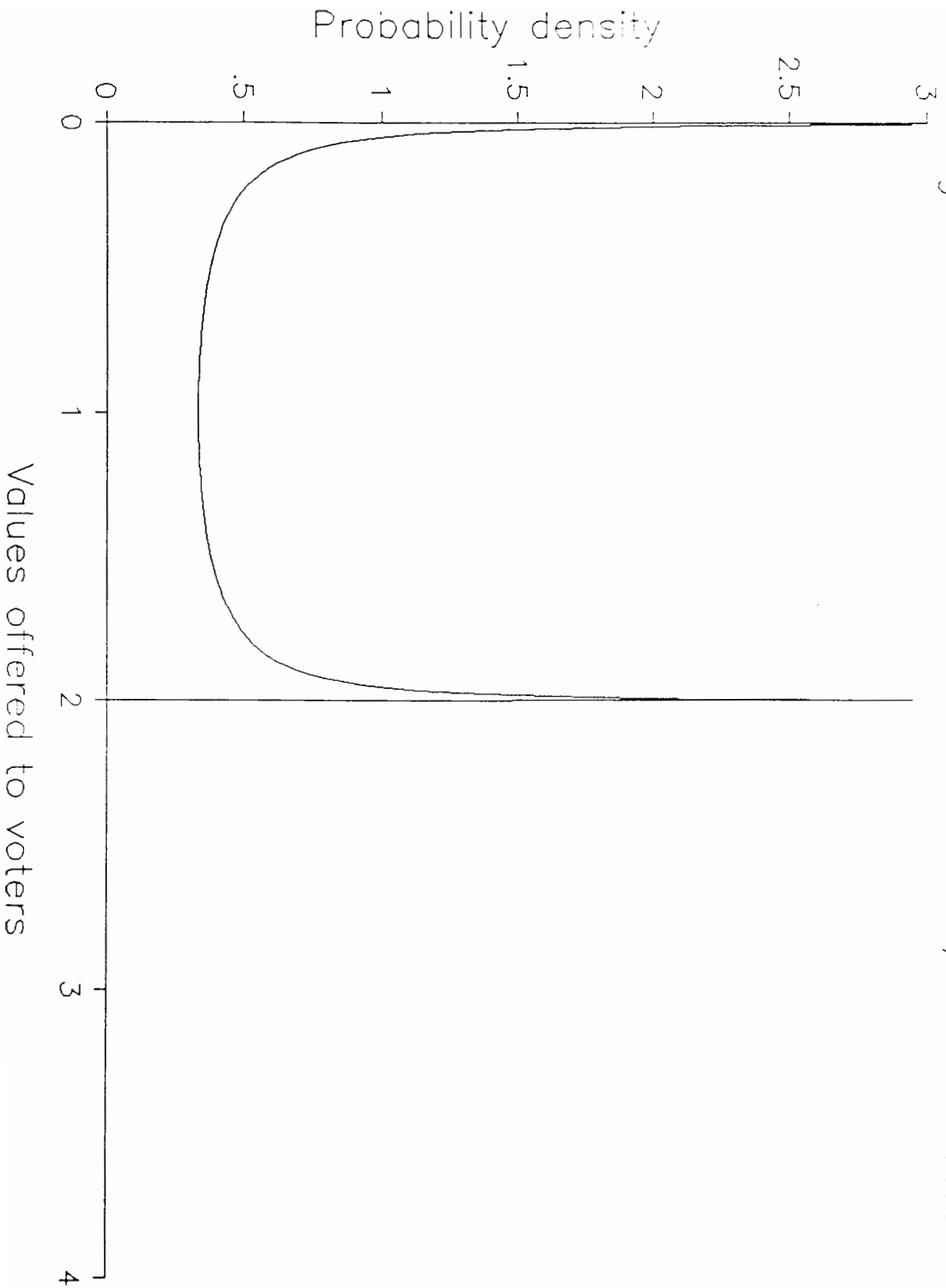


Figure 3. Offer distribution with 3 votes, 4 candidates.

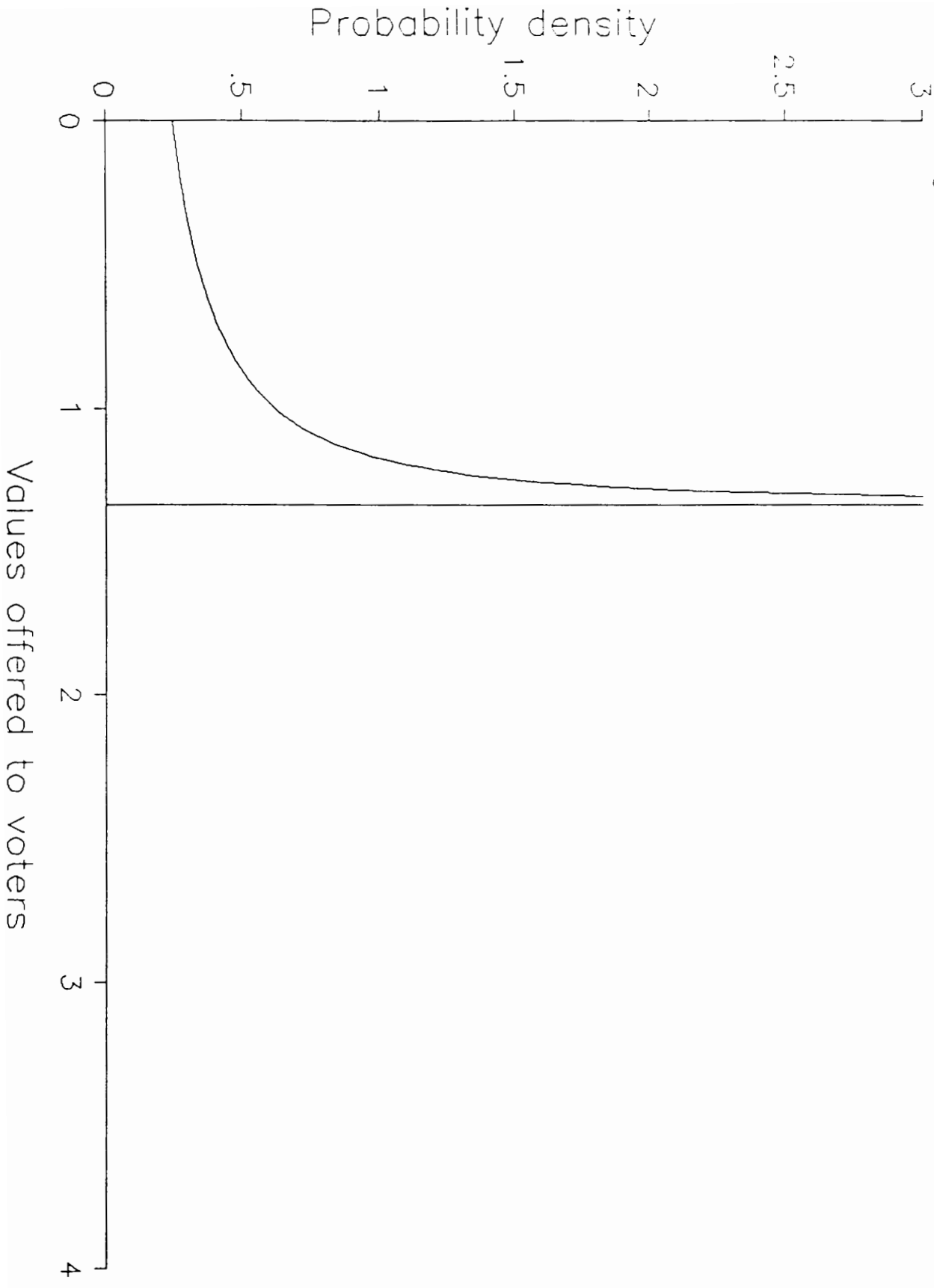


Figure 4. Offer distribution with ranking points (1,.5,.5,0).

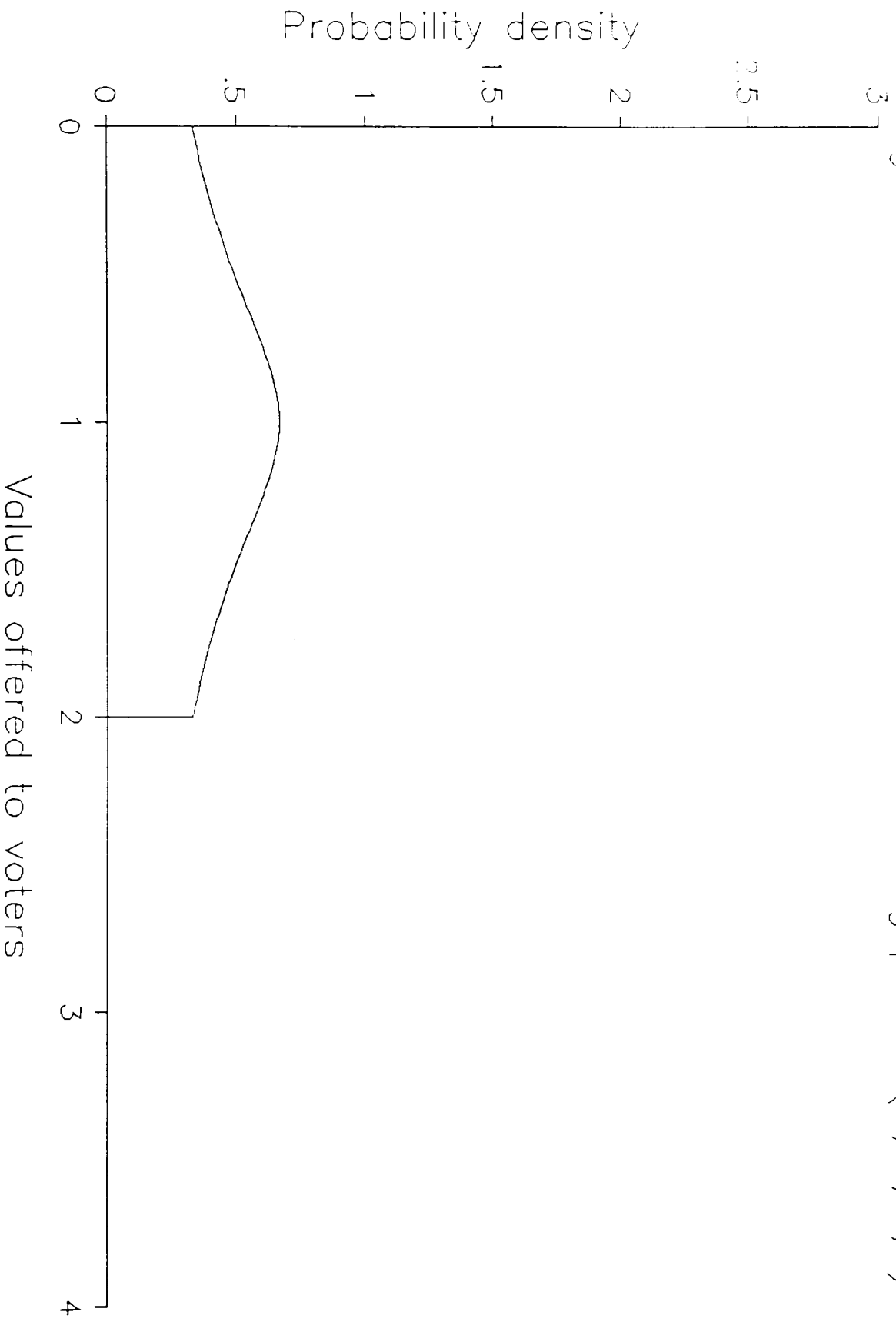


Figure 5. Offer distribution with approval voting, 4 candidates.

