Econometrica, Vol. 59, No. 1 (January, 1991), 1–23

AGGREGATION AND SOCIAL CHOICE: A MEAN VOTER THEOREM

BY ANDREW CAPLIN AND BARRY NALEBUFF[1]

A celebrated result of Black (1948a) demonstrates the existence of a simple-majority winner when preferences are single-peaked. The social choice follows the preferences of the median voter: the median voter's most-preferred outcome beats any alternative. However, this conclusion does not extend to elections in which candidates differ in more than one dimension. This paper provides a multi-dimensional analog of the median voter result. We provide conditions under which the *mean* voter's most preferred outcome is unbeatable according to a 64%-majority rule. The conditions supporting this result represent a significant generalization of Caplin and Nalebuff (1988).

The proof of our mean voter result uses a mathematical aggregation theorem due to Prékopa (1971, 1973) and Borell (1975). This theorem has broad applications in economics. An application to the distribution of income is described at the end of this paper; results on imperfect competition are presented in the companion paper, Caplin and Nalebuff (1991).

KEYWORDS: Condorcet paradox, log-concavity, mean voter, 64%-majority rule.

1. INTRODUCTION

A CELEBRATED RESULT of Black (1948a) demonstrates the existence of a simple-majority winner when preferences are single-peaked. The social choice follows the preferences of the median voter: the median voter's most-preferred outcome beats any alternative. However, this conclusion does not extend to an election with more than one issue at stake. This paper provides a multi-dimensional analog of the median voter result. We provide conditions under which the *mean* voter's most preferred outcome is unbeatable according to 64%-majority rule. The general conditions supporting this result represent a significant extension of Caplin and Nalebuff (1988).

Caplin and Nalebuff (1988) began our investigation into the positive properties of voting under a 64%-majority rule. When there is a sufficient similarity of voter preferences, we showed that there exists an unbeatable proposal, and furthermore no cycles are possible. Our notion of similarity of voter preferences relied on a concave density over voters' most-preferred proposals. More voters must favor intermediate positions than the average of those favoring extremes. This paper shows that the earlier results generalize to a much broader class of distributions. The requirement of a concave density is relaxed to allow for all log-concave densities. This includes such important distributions as the truncated normal, exponential, and Weibull. Our new results extend beyond the log-concave class. We provide a general bound on the minimum majority rule needed to support the mean voter's most preferred outcome. The bound depends on a concavity index for the distribution of preferences and the dimensionality of the issue space.

[1] We thank Jean-Michel Grandmont, James Heckman, Alvin Klevorick, and an anonymous referee for their comments and the National Science Foundation Grant #SES8909036 and Princeton University's John M. Olin Program for financial support.

The shift from median voter to mean voter requires a new mathematical approach. Black's median voter result is straightforward. When preferences are single-peaked, the issue at stake must be one-dimensional and proposals can be arranged in a linear order from left to right. Those favoring any proposal to the left of the median voter's favorite are outnumbered by the group consisting of the median voter and those to the right. In contrast, the demonstration of our mean voter theorem requires the use of a novel aggregation technique. This technique has broad applications in economics. An application to the distribution of income is outlined below. Results on imperfect competition are presented in the companion paper, Caplin and Nalebuff (1991).

Section 2 presents the aggregation theorems of Prékopa (1971, 1973) and Borell (1975). These results provide the mathematical foundation for our approach. Section 3 offers a brief summary of earlier results on super-majority rule. Section 4 presents our general assumptions, and provides examples of models that are covered by the assumptions. The main theorem is proved in Section 5, and discussed in detail in Section 6. Applications of the aggregation theorems to the Roy model of self-selection in the labor market follow in Section 7. Section 8 concludes.

2. THE PRÉKOPA-BORELL THEOREM

In this section we introduce the recent extensions of the Brunn-Minkowski theorem due to Prékopa (1971) and Borell (1975).[2] The theorem concerns inheritance of concavity properties under the integral sign. In order to discuss this aggregation property, we first introduce a general notion of concavity due to Avriel (1972). Following the general definition of ρ-concavity, we present a statement and interpretation of the Prékopa-Borell theorem.

DEFINITION: Consider $\rho \in [-\infty, \infty]$. For $\rho > 0$, a nonnegative function, f, with convex support $B \subset R^n$ is called *ρ-concave* if $\forall \alpha_0, \alpha_1 \in B$,

$$f(\alpha_\lambda) \geq \left[(1-\lambda) f(\alpha_0)^\rho + \lambda f(\alpha_1)^\rho\right]^{1/\rho}, \quad 0 \leq \lambda \leq 1, \tag{2.1}$$

where $\alpha_\lambda = (1-\lambda)\alpha_0 + \lambda\alpha_1$. For $\rho < 0$, the condition is exactly as above except when $f(\alpha_0)f(\alpha_1) = 0$, in which case there is no restriction other than $f(\alpha_\lambda) \geq 0$. Finally, the definition is extended to include $\rho = \infty, 0, -\infty$ through continuity arguments as discussed below.

For ρ positive, the definition states that f^ρ is concave while for ρ negative, $-f^\rho$ is concave.[3] Higher values of ρ correspond to more stringent variants of concavity; a ρ-concave function is also ρ'-concave for all $\rho' < \rho$. This follows as a

[2] Bonnesen and Fenchel (1987) provide a clear presentation of the Brunn-Minkowski theorem. Das Gupta (1980) and Dharmadhikari and Joag-Dev (1988) provide more accessible proofs of the later work and clarify its relation to the Brunn-Minkowski theorem.

[3] This assumes that f^ρ defines a function. If there are multiple solutions (such as when $f(\alpha) = \alpha^2$, $f^{1/2} = \pm\alpha$), then the statement applies to the unique positive root of f^ρ.

result of Hölder's Inequality, which shows that the right-hand side of equation (2.1) is monotonically increasing in ρ (see Hardy et al. (1934)).

At $\rho = \infty$, the right-hand side is set to its limiting value, $\max[f(\alpha_0), f(\alpha_1)]$. This condition can be satisfied only when f is uniform over its support.

$\rho = 1$ is the standard definition of concavity.

$\rho = 0$ corresponds to log-concavity of f. Here it may prove helpful to think of the right-hand side of (2.1) as a C.E.S. function with elasticity of substitution ρ. Using L'Hospital's rule, the C.E.S. approaches Cobb-Douglas as $\rho \to 0$. The condition (2.1) becomes $f(\alpha_\lambda) \geqslant f(\alpha_0)^{(1-\lambda)} f(\alpha_1)^{\lambda}$. Taking logarithms leads to the log-concavity condition,

$$\ln\left[f(\alpha_\lambda)\right] \geqslant (1-\lambda)\ln\left[f(\alpha_0)\right] + \lambda \ln\left[f(\alpha_1)\right].$$

$\rho = -\infty$ is the weakest condition. The right-hand side takes its limiting value of $\min[f(\alpha_0), f(\alpha_1)]$. The condition requires only that f be quasi-concave.

The definition of ρ-concavity is based on Hardy, Littlewood, and Polya's (1934) generalized mean function. We use this interpretation to further illustrate the condition, comparing the cases $\rho = 1$, $\rho = 0$, and $\rho = -1$. Let $f(0) = a$ and $f(1) = b$.

- $\rho = 1$: Concavity of f requires $f(\frac{1}{2}) \geq (a+b)/2$, the *arithmetic* mean of a and b.
- $\rho = 0$: Concavity of $\ln(f)$ requires $f(\frac{1}{2}) \geqslant \sqrt{ab}$, the *geometric* mean of a and b.
- $\rho = -1$: Convexity of $1/f$ requires $f(\frac{1}{2}) \geq 2ab/(a+b)$, the *harmonic* mean of a and b.

The use of ρ-concavity is new to the economics literature. Of special importance to our mean voter result is the class of log-concave densities, $\rho = 0$. This includes the multivariate beta, Dirichlet, exponential, gamma, Laplace, normal, uniform, Weibull and Wishart distributions. Below, we illustrate log-concavity for the normal and Weibull distributions. In some of the other cases, log-concavity requires restrictions on the parameter values; the restrictions and proofs for the beta, Dirichlet, and Wishart distributions are provided in Prékopa (1971). The argument for multivariate gamma, Laplace, and exponential distributions follows from the definitions in Johnson and Kotz (1972).

- The density of an n-dimensional normal distribution is:

$$f(\alpha) \propto e^{-1/2(\alpha-\mu)'\Sigma^{-1}(\alpha-\mu)}, \qquad \alpha \in R^n,$$

where Σ^{-1} is a positive definite matrix. Thus $f(\alpha)$ is log-concave as

$$(\alpha-\mu)'\Sigma^{-1}(\alpha-\mu)$$

is a convex function.

- The density of an n-dimensional Weibull distribution (also known as Type I extreme value or Gumbel distribution) is:

$$f(\alpha) = \prod_i e^{-e^{-\alpha_i}} e^{-\alpha_i}.$$

Each term in the product is log-concave in α. Hence the product is also log-concave. More generally, any multi-dimensional density which is the product of log-concave marginal densities is itself log-concave.

There are additional valuable insights on the case $\rho = 0$ provided in Prékopa (1971) and Brascamp and Lieb (1976). For example, they show that any convolution of log-concave densities is itself log-concave: indeed log-concavity is the weakest condition which ensures that convolutions are unimodal. For this reason, it is sometimes referred to as strong unimodality.

Our results on the mean voter theorem also apply to the class of distributions between 0 and $-1/(n+1)$ concave. These include the multivariate Cauchy, Pareto, F distributions, and t distributions. The specific ρ value and proofs for each of these cases follows from Borell (1975) and are exposited at greater length in Dharmadhikari and Joag-Dev (1988).

- The F distribution with degrees of freedom $(c_0, \ldots, c_n)$ has density:

$$f(\alpha) \propto \left[\prod_{i=1}^{n} \alpha_i^{1/2c_i - 1}\right]\left[c_0 + \sum_{i=1}^{n} c_i \alpha_i\right]^{-C/2},$$

with $\alpha_i, c_i > 0$ and $C = \Sigma_0^n c_i$. Here $\rho = -1/(n + (c_0/2))$.

- The density of the n-dimensional Pareto distribution is:

$$f(\alpha) \propto \left(\sum_{i=1}^{n} \alpha_i/\theta_i - n + 1\right)^{-(a+n)}, \qquad \alpha_i > \theta_i > 0, \quad a > 0.$$

This is $\rho = -1/(n+a)$ concave.

- The density of the n-dimensional Student's t distribution with a degrees of freedom is:

$$f(\alpha) \propto \left[1 + \frac{1}{a}(\alpha' - \eta') M^{-1}(\alpha - \eta)\right]^{-(a+n)/2}, \quad M^{-1} \text{ positive definite.}$$

This too is $\rho = -1/(n+a)$ concave since $f^{-1/(n+a)}$ is the square root of a quadratic form and hence convex. The case $a = 1$ corresponds to the multivariate Cauchy distribution, so that $\rho = -1/(n+1)$.

In many cases, economic reasoning requires that α be positive. A truncation of the density causes no additional difficulty. The same value of ρ applies to any truncations of the above distributions provided only that the support set is convex. This ability to handle truncated distributions is particularly important in the application of the normal and t distributions to the voting problem. Since these distributions are centrally symmetric, there would be a simple-majority winner in the absence of truncation. While symmetry is lost in truncation, our results continue to apply.

We are now ready to present a statement of the Prékopa-Borell theorem.

THEOREM (Prékopa-Borell): *Let f be a probability density function on R^n with convex support B. Take any measurable sets A_0 and A_1 in R^n with $A_0 \cap B \neq \varnothing$ and*

$A_1 \cap B \neq \varnothing$. *For* $0 \leqslant \lambda \leqslant 1$, *define* $A_\lambda = (1-\lambda)A_0 + \lambda A_1$, *the Minkowski average of the two sets.*[4]

If $f(\alpha)$ *is a* ρ*-concave function,* $\rho \geqslant -1/n$, *then*

$$\int_{A_\lambda} f(\alpha)\,d\alpha \geqslant \left[(1-\lambda)\left(\int_{A_0} f(\alpha)\,d\alpha\right)^{\rho/(1+n\rho)} + \lambda\left(\int_{A_1} f(\alpha)\,d\alpha\right)^{\rho/(1+n\rho)}\right]^{(1+n\rho)/\rho}.$$

As a first step in interpreting the theorem, it is helpful to parameterize the region of integration by λ and define the parameterized cumulative integral,

$$F(\lambda) \equiv \int_{A_\lambda} f(\alpha)\,d\alpha.$$

The theorem implies that ρ-concavity of f translates into $\rho/(1+n\rho)$-concavity of the cumulative integral in the parameter λ.

There are several different proofs of this general result due to Borell (1975), Brascamp and Lieb (1976), and Das Gupta (1980). One technique of proof is to first establish a more general result for $n = 1$. This argument is based on a sophisticated application of Hölder's inequality. The bound for higher dimensions then follows by induction. The integral over $A_\lambda \subset R^n$ is broken up into a double integral over R^{n-1} parameterized by the nth coordinate. By the inductive hypothesis, the integral over R^{n-1} is $\rho^* = \rho/(1+(n-1)\rho)$ concave. Integrating this over the nth coordinate and applying the $n = 1$ result shows that $F(\lambda)$ has concavity

$$\rho^*/(1+\rho^*) = \frac{\rho}{1+n\rho},$$

as stated by the theorem.

In one dimension, consider the example,

$$F(b) = \int_0^b f(\alpha)\,d\alpha.$$

By the Prékopa-Borell theorem, ρ-concavity of f implies $\rho/(1+\rho)$-concavity of $F(b)$. In this example, we can illustrate the theorem directly. Take $f(\alpha) = \alpha^m$ on $\alpha \geqslant 0$. Note that $f^{1/m}$ is linear so that f is ρ-concave for $\rho \leqslant 1/m$. Integration shows that $F(b) = b^{m+1}/(m+1)$, which is ρ-concave for $\rho \leqslant 1/(m+1)$.This is exactly the bound provided by the Prékopa-Borell theorem, $(1/m)/(1+1/m) = 1/(m+1)$.

More generally, the Prékopa-Borell theorem applies to integrals of ρ-concave functions over any parameterized regions A_b provided that A_{b_λ} contains

[4] For a given λ, $0 \leqslant \lambda \leqslant 1$, the Minkowski average A_λ is defined as all points of the form $x_\lambda = (1-\lambda)x_0 + \lambda x_1$, with $x_0 \in A_0$, $x_1 \in A_1$.

the λ-weighted Minkowski average of A_{b_0} and A_{b_1}. Parameterized integrals of this form arise frequently in economics. To see this, we illustrate the parameterizations for which the Prékopa-Borell theorem applies to problems in social choice, income distribution, and imperfect competition.

Define the function $F(b)$ by

$$F(b) \equiv \int_{\{\alpha \in R^n:\, W(\alpha) \leqslant b\}} f(\alpha)\, d\alpha.$$

When $W(\alpha) = \alpha \cdot \pi$, for some fixed gradient vector $\pi \in R^n$, then the regions of integration are half-spaces. It follows immediately that the half-space defined by b_λ equals the Minkowski average of the half-spaces defined by b_0 and b_1 so that the Prékopa-Borell theorem applies to the function $F(b)$. This is the form in which we apply the result to social choice. In fact, for any convex function $W(\alpha)$, integration using the lower contour sets satisfies the Minkowski averaging requirements: given $W(\alpha_0) \leqslant b_0$, $W(\alpha_1) \leqslant b_1$, and W convex it follows that α_λ is contained in the lower contour set for b_λ,

$$W(\alpha_\lambda) \leqslant (1-\lambda)W(\alpha_0) + \lambda W(\alpha_1) \leqslant (1-\lambda)b_0 + \lambda b_1 = b_\lambda.$$

This is the form in which the theorem is applied to study income distribution. An equivalent description of the regions of integration is $\{\alpha \in R^n:\ -W(\alpha) \geqslant -b\}$. Thus the results also apply to the upper contour sets of any concave function. This is the form in which the theorem is applied in our companion study of imperfect competition (Caplin and Nalebuff (1991)).

Das Gupta (1980) presents three different proofs of the Prékopa-Borell theorem, and provides a thorough discussion of the history of the theorem. The most important result is seen to be the Brunn-Minkowski theorem. In fact, the Brunn-Minkowski theorem is a special case of Prékopa-Borell where A_0 and A_1 are convex and the density is uniform. A uniform density corresponds to $\rho = \infty$, so that the value of $\rho/(1+n\rho)$ is taken to be $1/n$, the limit as $\rho \to \infty$. The parameterized family of convex sets for which the Minkowski averaging property holds is referred to as a concave family of convex sets. Since the density function is uniform, the integrals correspond to volumes. The Brunn-Minkowski theorem translates to the statement that the nth root of the volume of a concave family of convex sets in n-dimensions is a concave function of the integration parameter.

Our earlier work on 64%-majority rule (Caplin and Nalebuff (1988)) involved an indirect (and unconscious) application of the Brunn-Minkowski theorem. We used a theorem due to Grunbaum (1960) and Hammer (1960) that involved performing a "Schwarz symmetrization" on a convex set, an operation which symmetrizes a convex set. A central property of this operation is that it preserves convexity: it turns out that this fact is equivalent to the Brunn-Minkowski theorem (see Bonessen and Fenchel (1987)).

The connection to the Brunn-Minkowski theorem highlights the geometric content of the result. Section 6 provides a brief discussion of the geometric intuition when explaining the super-majority rule bound. The geometric inter-

pretation is developed more systematically in our companion paper on imperfect competition.

The paper by Prékopa (1971) establishes the inheritance property for the central case of log-concavity in one-dimension ($\rho = 0$, $n = 1$). This would be an easy result if we were multiplying log-concave functions; what makes the result challenging is that it demonstrates inheritance of log-concavity under a process of addition. The one-dimensional result can also be proved using a completely different mathematical technique based on the theory of total positivity, exposited by Karlin (1968). Papers by Pratt (1981), Flinn and Heckman (1982), and Jewitt (1987, 1988) apply the one-dimensional inheritance of log-concavity to problems in statistical theory, search theory, and financial economics, respectively.

Prékopa (1973) generalizes the inheritance of log-concavity to integrals in any number of dimensions ($n \geqslant 1$). The inheritance property of ρ-concavity for $\rho < 0$, both in one and in higher dimensions, is due to Borell (1975) (see also Rinott (1976) and Brascamp and Lieb (1976)). We apply the full multi-dimensional power of this aggregation result to characterize properties of voting under super-majority rule.

3. THE VOTING PROBLEM

There are three serious problems associated with voting by simple majority rule. There is the Condorcet paradox: for any given proposal there is an alternative preferred by a majority of the population (Condorcet (1785)). The nonexistence of a Condorcet winner is generic (Plott (1967) and Rubinstein (1979)). Given nonexistence of a Condorcet winner, it is always possible to find a sequence of majority victories leading from any given outcome to any other (McKelvey (1979)): thus, an individual in control of the voting agenda has great power over the final outcome.

Against this, the primary positive result on voting by simple majority rule is due to Black (1948a). Black showed that when preferences are single-peaked, there always exists a simple majority winner, the proposal most favored by the median voter. Unfortunately, the condition of single-peakedness imposes strong restrictions on preferences, and only applies to issues which are one-dimensional.

Black (1948b) also initiated the study of voting under super-majority rules. When more than 50% must vote against the status quo to overturn it, this increases the possibilities for finding an unbeatable proposal and reduces the scope for voting cycles. However, Greenberg (1979) demonstrated that with an n-dimensional issue space, the *minimum* majority size needed to ensure the existence of an unbeatable proposal is $n/(n+1)$. Without any prior bound on n, only a unanimity rule guarantees existence of an unbeatable proposal.

Caplin and Nalebuff (1988) provide a more positive result on voting under super-majority rules. When individual preferences are Euclidean and the most preferred proposals are distributed according to a concave probability density,

the majority size needed to avoid cycles and ensure existence of an unbeatable proposal is no higher than $1-[n/(n+1)]^n$. Our earlier paper derives its title from the fact that $1-[n/(n+1)]^n$ is increasing in n, with limit $1+(1/e)$, or just under 64%. This paper applies the Prékopa-Borell theorem to provide a considerable extension and simplification of our earlier results. In particular, for all log-concave densities, the proposal most preferred by the mean voter is unbeatable under 64%-majority rule.

4. THE MODEL

There is a social decision to be made. The compact set of proposals among which society can choose is denoted by X. Elements of X are represented as vectors in w-dimensional Euclidean space. Preferences vary across society as summarized by a vector $\alpha \in R^n$, an n-dimensional index of types. The preferences of an α-type for proposals $x \in X$ are represented by a continuous utility function $U(\alpha, x)$. The distribution of types across society is represented by a probability measure with density f on utility parameters $\alpha \in R^n$. We refer to the *mean* voter as the type $\bar{\alpha}$ lying at the center of gravity of the probability distribution over α.[5] Correspondingly, the mean voter's most preferred outcome in X is denoted by $\bar{x}$.

We are interested in showing the conditions under which $\bar{x}$ is a δ-majority rule winner. A δ-majority winner is a proposal in X which is preferred by at least a fraction $(1-\delta)$ of the population to any alternative in pairwise comparisons. Formally, let $m(\bar{x}, y)$ denote the proportion of the population that strictly prefers y to $\bar{x}$, and $m(\bar{x}) = \sup_{y \in X} m(\bar{x}, y)$ denote the maximal vote against $\bar{x}$.[6] Then $\bar{x}$ is a δ-majority winner if and only if $\delta \geqslant m(\bar{x})$.

The fundamental result of this paper is Theorem 1, which shows how to place a bound on $m(\bar{x})$ as a function of the degree of concavity and the dimensionality of preferences. The result requires restrictions both on the form of the utility functions and the distribution of types across society.

ASSUMPTION A1 (Linear Preferences): *Preferences can be represented in a linear form*:

$$U(\alpha, x) = \sum_{k=1}^{n} \alpha_k t_k(x) + t_{n+1}(x) \tag{4.1}$$

where U: $R^n \times X \mapsto R$, *and the functions* $t_k(x)$ $X \to R$, *for* $1 \leqslant k \leqslant n+1$.

[5] The mean voter is defined by the type

$$\bar{\alpha} = (\bar{\alpha}_1, \ldots, \bar{\alpha}_n), \quad \text{where} \quad \bar{\alpha}_i = \int_{\alpha \in R^n} \alpha_i f(\alpha)\, d\alpha.$$

[6] We consider hyperdiffuse distributions, so that voter indifference can be ignored.

Assumption A1 is due to Grandmont (1978).[7] The restriction implies a separability of issues in determining voter preferences. Each voter evaluates a platform by a weighted sum of the utility from the position in each dimension. While the weights may differ, the utility valuations are common across the population. The distribution of weights is described by a density function, $f(\alpha)$, on $\alpha \in R^n$. We require that there be some measure of concavity in the distribution of these weights. Caplin and Nalebuff (1988) required $\rho \geqslant 1$. Here we relax this consensus condition to include the general class of ρ-concave densities.

ASSUMPTION $A2_\rho$ (ρ-concavity): *The probability density of consumers' utility parameters is a ρ-concave function over its support, B, which is a convex subset of R^n with positive volume.*

Note that A1 and $A2_\rho$ are joint assumptions: all that matters is that there exists some specification of preferences such that A1 and $A2_\rho$ are both satisfied.

The fundamental result is the following theorem, which places a bound ensuring the existence of a δ-majority winner. The bound is a function of both the dimensionality of preferences, n, and the concavity, ρ, of the density.[8] We call this result a mean voter theorem as the bound is based on the proposal $\bar{x} \in X$ most preferred by the voter whose preferences lie at the center of gravity of the population distribution.

THEOREM 1: *When voter preferences $\alpha \in R^n$ satisfy A1 and $A2_\rho$, then for $\rho \geqslant -1/(n+1)$, $\bar{x}$ is a δ-majority rule winner when the majority size equals*

$$d(n,\rho) \equiv 1 - \left[\frac{(n+1/\rho)}{(n+1+1/\rho)}\right]^{n+1/\rho}.$$

The proof of the theorem is based on an application of the Prékopa-Borell aggregation theorem and is presented in Section 5. The interpretation of the bound $d(n,\rho)$ is the focus of Section 6. Our discussion of the theorem begins with a series of examples that illustrate the applicability of A1. In each case, we discuss the meaning of $A2_\rho$ in the context of the example.

• Euclidean Preferences: Each person votes for the proposal closest to their most-preferred point. We identify the type by the most-preferred point,

$$U(\alpha, x) = -\|x - \alpha\|.$$

When written in this form, it is not directly apparent that α is the weighting

[7] Assumption A1 may be generalized to include Grandmont's class of intermediate preferences. Because these preferences are not necessarily transitive, we would then have to add an assumption that the mean voter has a most preferred point in X.

[8] While n refers to the dimension of the preference parameter, typically α and X will be the same dimensionality ($n = w$). In these cases, we may also interpret the bound as depending on the dimensionality of the issue space.

function for the different issues. However, these preferences are equivalent to

$$U(\alpha, x) = -\|x-\alpha\|^2 = -[x\cdot x - 2\alpha\cdot x + \alpha\cdot\alpha],$$

which is in the linear form once the irrelevant term $(\alpha\cdot\alpha)$ is removed.[9] The meaning of A2 in this example is that there is ρ-concavity or some weak form of consensus in the distribution of most-preferred points.

• Linear Preferences: Each voter evaluates a platform by a weighted sum of the positions,

$$U(\alpha, x) = \alpha\cdot x.$$

This includes the important cases of logit and probit models discussed in Example 4.1 below. With $t_k(x) = x_k$, each individual can rank the issues by their importance using the weights α_k. The meaning of A2 is that there is some form of consensus over the rankings of issues; in particular, the population is not split into groups that each place a very high weight on their "pet" issue but regard other issues as irrelevant.

• C.E.S. Preferences: Each voter evaluates a platform by a weighted sum of the utility from the position in each dimension:

$$U(\alpha, x) = \sum_{k=1}^{n} \alpha_k x_k^{\rho}.$$

In the linear case, the rate of substitution between issues was independent of the position. Here, with $\rho < 1$, there is a diminishing marginal rate of substitution. One special case of C.E.S. is the Cobb-Douglas utility function which arises as $\rho \to 0$,

$$U(\alpha, x) = \sum_{k=1}^{n} \alpha_k \ln x_k.$$

Two particularly important and testable cases that satisfy both A1 and $A2_0$ are the logit and probit models.

EXAMPLE 4.1: The linear utility function includes the standard qualitative response or discrete choice models used in the econometric literature, such as multinomial logit, multinomial probit, and the random coefficients models (see McFadden (1981)). To apply the discrete choice models to voting, we first divide the set of issues in a candidate's political platform into two classes consisting of observables and unobservables respectively. The observables may be thought of as the candidate's public positions. The unobservables may correspond to the voters' perceptions of less quantifiable characteristics (such as charisma and honesty). Additionally, the unobservable characteristics may represent voter uncertainty as to the candidate's exact position on any particular issue.

[9] Note that the utility function may be quadratic in distance, linear, or any other power of distance; in all cases, the utility function represents the same preferences.

The utility function for consumer of type α is

$$U(\alpha, x) = \beta \cdot x^0 + \alpha \cdot x^u,$$

where x^0 are the observable characteristics of the candidate, x^u are the unobservable characteristics. The logit and probit models further specialize to the case where β is common across all individuals, and there is a single idiosyncratic unobservable characteristic associated with each candidate. Hence the utility value for candidate i reduces to $\beta \cdot x^0 + \alpha_i$. Logit and probit involve distributional assumptions concerning $f(\alpha)$. Logit uses the multi-dimensional independent Weibull while probit uses the normal; both are covered by $A2_0$. The random coefficients model extends probit to allow for the possibility that β is normally distributed across the population; given that β enters the utility function linearly, this too is included in $A2_0$.

In applications, an advantage of the logit formulation is that the vote for any position against any alternative is readily computed as a function of only the observable characteristics. The vote for position i, $i = [0, 1]$, is

$$\frac{e^{\beta \cdot x_i^0}}{e^{\beta \cdot x_0^0} + e^{\beta \cdot x_1^0}}.$$

Our second example is drawn from the literature on financial decision-making with incomplete markets. Here agents are valuing corporate profits rather than party platforms.

EXAMPLE 4.2: Without complete markets, we do not expect shareholder unanimity over what constitutes the optimal production plan. Each shareholder values the potential payoffs of a firm differently. An individual of type α values a firm at $\alpha \cdot x$, where x is the vector of profits in the different states and α is the individual's shadow price for money in the different states.

In terms of A2, the natural way to measure the distribution of types is by the total size of their shareholdings. In exactly this context, Drèze (1974) suggested that firms adopt the plan that maximizes the valuation of the mean shareholder (the person whose shadow prices lie at the center of gravity of the distribution of shadow prices, where the distribution reflects the number of shares held by each type). Adoption of the mean shareholder's most preferred plan brings us directly under Theorem 1. It follows that this "pseudo-value" maximizing plan will be unbeatable under a $d(n, \rho)$-majority rule, where n is the dimensionality of the state space. The restriction to ρ-concavity is a statement about the dispersion of valuations across the shareholders of a particular company.[10]

In general terms, A1 and A2 are to be seen as imposing a family of distinct forms of consensus on social preferences.

[10] The self-selection of investors into different stocks may have interesting implications for the degree of consensus among the shareholders of a given firm. Individuals will hold stocks with payoffs that are positively correlated with their preference vector, α. As a result, the part of the population holding stock i may be more homogeneous in their distribution of α than the population as a whole.

5. PROOF OF THE MAIN THEOREM

The strategy of our proof is to show how ρ-concavity of the population density translates into a concavity property for a cumulative distribution. With A1, the division of the population between distinct proposals is always defined by a hyperplane in the parameter space. We show that the problem of bounding the maximum vote against $\bar{x}$ reduces to bounding the population contained in any half-space passing through the center of gravity. The population in any of these half-spaces is then bounded using the concavity index for the density function.

To begin the argument, we introduce a distribution function for the population contained in the intersection of the parameter space and a family of half-spaces. Given a vector π in R^n, define the cumulative distribution function $F_\pi(b)$ by

$$F_\pi(b) \equiv \int_{\{\alpha \in R^n:\ \alpha \cdot \pi \leqslant b\}} f(\alpha)\, d\alpha. \tag{5.1}$$

Here $F_\pi(b)$ is the proportion of the population with $\alpha \cdot \pi \leqslant b$. It follows from the nonatomic nature of the population distribution under A2 that $F_\pi(b)$ is continuous in b. But our results require far more than continuity. As confirmed in the discussion in Section 2, the Prékopa-Borell theorem applies to the function $F_\pi(b)$ since the set $\{\alpha \in R^n:\ \alpha \cdot \pi \leqslant b_\lambda\}$ is the Minkowski average of the sets $\{\alpha \in R^n:\ \alpha \cdot \pi \leqslant b_0\}$ and $\{\alpha \in R^n:\ \alpha \cdot \pi \leqslant b_1\}$. Application of the Prékopa-Borell theorem immediately yields Proposition 1, which shows how the parameter ρ influences the shape of the cumulative distributions $F_\pi(b)$.

PROPOSITION 1: *Under A1 and A2$_\rho$, the function $F_\pi(b)$ is $\rho/(1+n\rho)$-concave.*

This ρ-concavity property of F enables us to place a bound on the maximum vote against $\bar{x}$. Proposition 2 establishes this result by showing that the vote for $\bar{x}$ is always greater than the minimal value of $F_\pi(\bar{b}_\pi)$ across gradient vectors π, where $\bar{b}_\pi$ is the mean of b under $F_\pi(b)$,

$$\bar{b}_\pi = \int_R b\, dF_\pi(b).$$

PROPOSITION 2: *Under A1,*

$$1 - m(\bar{x}) \geqslant \min_\pi F_\pi(\bar{b}_\pi).$$

PROOF: With A1, the types indifferent between two proposals define a hyperplane in the space of utility parameters. In comparing proposals to $\bar{x}$, we know that the individual of type $\bar{\alpha}$ is always on the side of the hyperplane favoring $\bar{x}$. The reason is that $\bar{x}$ is $\bar{\alpha}$'s most preferred proposal, so that $\bar{\alpha}$ prefers it to everything else. Hence the minimal vote for $\bar{x}$ is bounded below by the minimum proportion of the population of B contained in any hyperplane passing through the centroid of the distribution, $\bar{\alpha}$. If we allow all possible

gradient vectors, then we can restrict our attention to the population below the centroid, $F_\pi(\bar{\alpha}\cdot\pi)$, without missing any of the populations in which we are interested,

$$(5.2)\qquad 1-m(\bar{x})\geqslant \min_{\pi} F_\pi(\bar{\alpha}\cdot\pi).$$

Finally, since the expectation is a linear operator, the value $\bar{\alpha}\cdot\pi$ is equal to $\bar{b}_\pi$, the expected value of b under $F_\pi(b)$. Substituting this into the lower bound in (5.2) establishes the proposition. *Q.E.D.*

Consider now an arbitrary ρ^*-concave distribution F with mean $\bar{b}$. We show that it is possible to provide a lower bound on $F(\bar{b})$, where the bound depends only on ρ^*. The intuition for this result is easiest for the case when F is concave. For all concave distributions, the mean is greater than the median. Thus $F(\bar{b})\geqslant 1/2$ with equality only when F is linear. When F is linear, this corresponds to an underlying uniform density; the mean equals the median and $F(\bar{b})=1/2$.[11] A similar argument allows us to place a lower bound on $F(\bar{b})$ for all ρ-concave functions.

PROPOSITION 3: *Given a cumulative distribution function $F(b)$ that is continuous and ρ^*-concave, the population below the mean satisfies*

$$(5.3)\qquad F(\bar{b})\geqslant\left[\frac{1}{1+\rho^*}\right]^{1/\rho^*}.$$

PROOF: We first establish that this bound in (5.3) is exact for the class of ρ^*-linear distributions introduced below. We then use these cases to verify the inequality for all other ρ^*-concave c.d.fs.

The following distributions are referred to as ρ^*-linear.

- $\rho^*>0$: $F(b)=b^{1/\rho^*}$, $b\in[0,1]$.
- $\rho^*=0$: $F(b)=e^b$, $b\leqslant 0$.
- $\rho^*<0$: $F(b)=(1-b)^{1/\rho^*}$, $b\leqslant 0$.

We also include as ρ^*-linear all distributions derived from the above by affine transforms of the domain.

In the first case, $\rho^*>0$ and $\bar{b}=1/(1+\rho^*)$. In the second, $\bar{b}=-1$ so that $F(\bar{b})=1/e$. In the third case, $1-\bar{b}=1/(1+\rho^*)$. Thus in all three cases,

$$F(\bar{b})=\left[\frac{1}{1+\rho^*}\right]^{1/\rho^*},$$

and (5.3) holds as an equality, as claimed.[12]

[11] Note that ρ-concavity has no predictive power for placing an *upper* bound on the population below the mean: it is possible that $F(\bar{b})$ may be arbitrarily close to 1. For example, $F(b)=kb$ for $b\in[0,1/(k+1)]$ and $F(b)=(k-1+b)/k$ for $b\in[1/(k+1),1]$ is a concave cumulative density. But in this case, $\bar{b}=1/(1+k)$ and $F(\bar{b})=k/(k+1)$, which may be arbitrarily close to 1.

[12] The limiting value as ρ^* approaches zero is $1/e$.

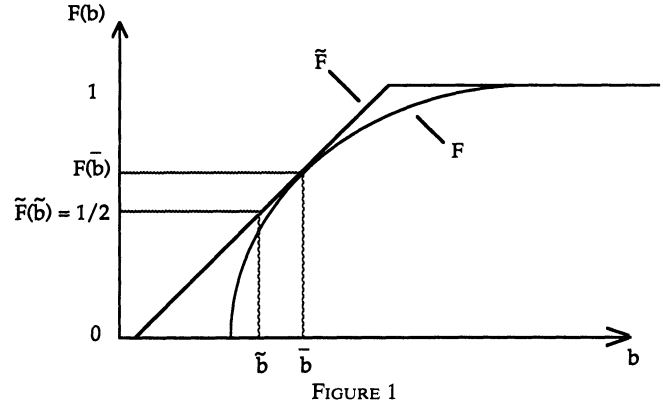

FIGURE 1

The proof that (5.3) holds for $F(b)$ strictly ρ^*-concave is separated into three parts according to the sign of ρ^*. In all three cases, the idea is to linearize the distribution around the point $\bar{b}$, resulting in a ρ^*-linear distribution denoted by $\tilde{F}(b)$. We then show that when each distribution is evaluated at its own mean, $\tilde{b}$ for $\tilde{F}$ and $\bar{b}$ for F, $F(\bar{b})$ is greater than $\tilde{F}(\bar{b})$. Combining this with the observation that (5.3) is satisfied as an equality for $\tilde{F}$ completes the proof.

First, consider cases with $\rho^* > 0$. We derive $\tilde{F}$ from F using a support of the concave function $F(b)^{1/\rho^*}$ at the point $\bar{b}$ (the mean of the distribution F). Formally, since $F(b)^{1/\rho^*}$ is concave, it lies everywhere below any supporting line at the point $\bar{b}$:

$$(5.4) \qquad F(b)^{1/\rho^*} \leqslant F(\bar{b})^{1/\rho^*} + k[b - \bar{b}] = \tilde{F}(b)^{1/\rho^*},$$

where k is a subgradient of $F(b)^{1/\rho^*}$ at $\bar{b}$. To turn $\tilde{F}(b)$ into a cumulative distribution, first note that k is strictly positive.[13] Thus the function $\tilde{F}(b)$ is strictly increasing, and setting the support $[c, d]$ so that $\tilde{F}(c) = 0$ and $\tilde{F}(d) = 1$ yields a new c.d.f. Note also that when we shift the support to $c = 0$ and $d = 1$, we see that $\tilde{F}(b)$ is ρ^*-linear. Hence the value of $\tilde{F}(b)$ when evaluated at its own mean, $\tilde{b}$, is $[1/(1+\rho^*)]^{1/\rho^*}$. Since the function $\tilde{F}(b)$ lies everywhere above $F(b)$, the expectations satisfy $\tilde{b} \leqslant \bar{b}$. Hence,

$$(5.5) \qquad F(\bar{b}) = \tilde{F}(\bar{b}) \geqslant \tilde{F}(\tilde{b}) = \left[\frac{1}{1+\rho^*}\right]^{1/\rho^*},$$

establishing (5.3). This argument is illustrated in Figure 1 above for the case of $\rho = 1$, so that F is concave and $\tilde{F}(\tilde{b}) = 1/2$.

When $\rho^* = 0$, the above proof applies with the following amendments. We use the function $\ln[F]$ in place of F^{1/ρ^*} to define $\ln \tilde{F}$ in equation (5.4). To turn this into a c.d.f. we restrict the argument b to be below the value d, where

[13] If $k = 0$, this implies $F(\bar{b}) = 1$, which in turn implies that $b = \bar{b}$ a.s., contradicting continuity of $F(b)$.

$\ln[\tilde{F}(d)] = 0$. Once again we arrive at a ρ^*-linear distribution, so that the inequalities in (5.5) are still valid, and (5.3) again holds.

Finally, we turn to cases with $\rho^* < 0$. Here ρ^*-concavity implies that F^{1/ρ^*} is a convex function. Hence we reverse (5.4) to define the new function $\tilde{F}^{1/\rho^*}$ which is tangent to F^{1/ρ^*} at $\bar{b}$, and lies everywhere *below* F^{1/ρ^*}. Once again, we pick the appropriate range to turn $\tilde{F}$ into a c.d.f.. Since $\rho^* < 0$, it remains true that $\tilde{F}$ lies everywhere above F, so that (5.5) follows as before, completing the demonstration of (5.3). *Q.E.D.*

The proof of Theorem 1 involves combining Propositions 1 to 3.

THEOREM 1: *When voter preferences* $\alpha \in R^n$ *satisfy* $A1$ *and* $A2_\rho$, *then for* $\rho \geqslant -1/(n+1)$, $\bar{x}$ *is a* δ*-majority rule winner when the majority size equals*

$$(5.6) \qquad d(n,\rho) \equiv 1 - \left[\frac{(n+1/\rho)}{(n+1+1/\rho)}\right]^{n+1/\rho}.$$

PROOF: By Proposition 1, A1 and $A2_\rho$ together imply that all functions F_π are ρ^*-concave,

$$(5.7) \qquad \rho^* = \frac{\rho}{1+n\rho}.$$

Hence by Proposition 3, the inequality (5.3) applies to all distributions F_π. Substituting (5.7) into (5.3) yields

$$\min_\pi F_\pi(\bar{b}_\pi) \geqslant \left[\frac{1}{1+\rho^*}\right]^{1/\rho^*} = \left[\frac{n+1/\rho}{n+1+1/\rho}\right]^{n+1/\rho} = 1 - d(n,\rho).$$

Proposition 2 establishes the theorem as $m(\bar{x}) \leqslant 1 - \min_\pi F_\pi(\bar{b}_\pi) \leqslant d(n,\rho)$. *Q.E.D.*

6. INTERPRETATION OF THEOREM 1

We now provide a detailed analysis of the bound of Theorem 1,

$$d(n,\rho) = 1 - \left[\frac{n+1/\rho}{n+1+1/\rho}\right]^{n+1/\rho}.$$

We first consider uniform densities over convex sets in R^n, which correspond to $\rho = \infty$. As ρ increases to infinity, so $d(n,\rho)$ increases to the limit,

$$d(n,\infty) = 1 - [n/(n+1)]^n.$$

This is the bound in Caplin and Nalebuff (1988). Observe that $[n/(n+1)]^n$ diminishes to a limit of $1/e$ as n rises, so that $d(n,\rho)$ increases to $1 - 1/e$, or just above 63%.

At $\rho = 1$,

$$d(n,1) = 1 - \left[\frac{(n+1)}{(n+2)}\right]^{n+1},$$

which corresponds to the uniform bound in one higher dimension, so that $d(n,1)$ also increases with n to $1 - 1/e$. Our original paper took its title from the observation that $1 - 1/e$ provides an upper-bound on the maximum vote against $\bar{x}$ for all concave densities, regardless of dimension.

But the general form of Theorem 1 allows us to extend this upper bound well beyond the concave case. For example, with $\rho = 1/m$,

$$d(n,1/m) = 1 - [(n+m)/(n+m+1)]^{(n+m)}.$$

This is the bound for the uniform distribution in $(n+m)$-dimensions. Once again, this remains bounded above by $1 - 1/e$.

With $\rho = 0$, $d(n,0)$ is interpreted as the limit of the expression as ρ shrinks to zero:

$$d(n,0) = 1 - 1/e,$$

regardless of dimension. Hence the upper bound of 64% applies even for distributions which are log-concave. We regard this as a very significant extension of the original result; the class of log-concave densities includes the beta, χ^2, Dirichlet, exponential, gamma, Laplace, normal, uniform, Weibull, and Wishart distributions, where only the uniform is also concave.

There is a geometric approach which provides additional insight into Theorem 1 for all cases with $\rho \geq 0$. When $\rho = 1/m$, imagine that we add m extra dimensions above the support set B, each representing the concave function $f^{1/m}$. This leads to a convex set with uniform density in $(n+m)$-space. The population in any division of B by a hyperplane is measured by the volume of this $(n+m)$-dimensional set in the appropriate half-space. Because this artificial addition of dimensions preserves all the values $F_\pi(b)$, the bound for the case with $f^{1/m}$ concave in n-dimensions is the same as the bound for f uniform in $n+m$ dimensions.

The simplest example is the case of a concave density over α in R^n. This may be viewed equivalently as a uniform density over an $n+1$-dimensional convex set where the density is represented as the height of the set. The transformation is more complicated when the density is not concave. We illustrate the equivalence for $f(\alpha) = \alpha^2$ in Figure 2 below. Although f is not concave, $f^{1/2}$ is concave: $\rho = 1/2$. This suggests that we should represent the density using 2

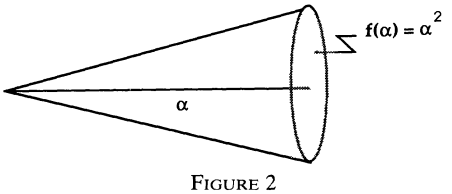

FIGURE 2

dimensions, as the cross-section volume of a cone with radius $f^{1/2} = \alpha$. The resulting figure is now a convex set in 3 dimensions.

The case $\rho = 0$ may be viewed as the limiting case of ρ strictly positive. In terms of derivatives, $\ln[f(\alpha)]$ strictly concave corresponds to $f''f - f'^2 < 0$, while $f^{1/m}$ concave corresponds to $f''f - f'^2 \leqslant -f'^2/m$. Thus with f twice continuously differentiable over a bounded support, strict log-concavity implies that $f^{1/m}$ is concave for some positive m, so that the upper bound of $1 - 1/e$ remains valid.

The formula takes on a different nature when $\rho < 0$. For example, the Student's t distribution with a degrees of freedom has $\rho = -1/(n + a)$. For $a \geqslant 1$,

$$d[n, -1/(n+a)] = 1 - [(a-1)/a]^a.$$

This number is always *above* $1 - 1/e$. However, $d[n, -1/(n+a)]$ falls to $1 - 1/e$ as a increases towards infinity. The 64% bound on the δ-majority remains valid even as ρ approaches zero from below.

It is important to emphasize this bound on the majority rule depends only on the degrees of freedom, a, and not the dimension of preferences, n. The same holds true for the F distribution and Pareto distribution. Here too $\rho = -1/(n + a)$ (although a has different interpretations). Thus in higher dimensions, the increase in ρ is just offset by the effect of n in the formulation of $d(n,\rho)$.

The bound of Theorem 1 is the best available. There exist ρ-concave distributions for which the maximum vote against $\bar{x}$ equals $d(n, \rho)$. This is easily seen for the case with Euclidean preferences and $\rho = \infty$. Consider a population uniformly distributed over the n-dimensional unit simplex. Here the center of gravity secures exactly $[n/(n+1)]^n$ of the vote against alternatives which approach it along the perpendicular to any face. For $\rho < \infty$, we can show the bound is tight using ρ-linear densities, as defined in the previous section.

The minimum majority size needed to ensure the existence of a δ-majority rule winner is called the Simpson-Kramer min-max majority, and the corresponding unbeatable proposal is the min-max point.[14] The proposal $\bar{x}$ need not be the min-max point and $d(n, \rho)$ is only an upper bound on the min-max majority. There are two distinct ways in which the bound on the min-max majority can be lowered from $d(n, \rho)$: by reducing dimension through a fixed point argument, and by moving away from the mean voter's choice.

In our earlier paper, we used a fixed point argument to reduce the bound on the min-max majority by transforming the case with f concave into a problem with f uniform on a convex set. Hence the uniform bound of $1 - [n/(n+1)]^n$ applies equally to the concave case. In a similar manner, the bound for the case

[14] There is an interesting parallel in the statistical literature where the min-max point is used to measure the "depth" of a set and is viewed as a multi-dimensional analog of the median. This interpretation is due to Tukey (1977) and has been developed by Huber (1985) and Donoho and Gasko (1987). Most recently, Nolan (1989) derived the distribution of the sample min-max point as an estimator for the true min-max point for spherically symmetric distributions.

$\rho = 1/m$ can be lowered to $1 - [(n + m - 1)/(n + m)]^{(n+m-1)}$, the bound for the uniform case in $(n + m - 1)$-dimensions. Unfortunately, we are no longer able to identify the min-max point. Once n is large, this reduction in dimension results only in a trivial reduction in the bound.

For cases with $\rho < 0$, $d(n, \rho)$ may be a less useful bound on the min-max majority. For example, as a falls to 1, $d[n, -1/(n + a)]$ rises to 1. But Greenberg (1979) establishes a general upper bound on the min-max majority of $(n - 1)/n$ which doesn't depend on either A1 or A2. Why does our bound ever rise above this upper limit?

Our argument is based on the mean voter's most preferred outcome as a proxy for the true min-max point. While the two points are the same for uniform distributions on the simplex, in other cases they may differ substantially. Thus the bound provided for the proportion of the volume of a body on either side of the centroid need not bear any close relation to the volume on either side of the min-max point. As the concavity restriction is relaxed, the centroid and the true min-max point may move further apart.

This observation also sheds light on the nature of Theorem 1. The underlying mathematical result places bounds on the population of regions defined by hyperplanes passing through the center of gravity of a distribution. These bounds, which depend on the degree of concavity of the associated density, also provide bounds on the min-max majority. For positive values of ρ, we lose practically nothing by replacing the true min-max point with the centroid of the distribution. Hence there is the additional implication that the centroid is a good approximation to the min-max point for densities with ρ positive.[15] This fact may turn out to be useful for computational purposes.

7. THE MEAN VOTER THEOREM APPLIED TO THE DISTRIBUTION OF INCOME

The mean voter theorem is based on a characterization of how the population of voters is distributed around its mean. In this section, we take a similar approach to show how the distribution of worker skills bounds the position of the mean in the overall distribution of income. This bound is independent of the wages; it depends only on the ρ-concavity index for the distribution of worker skills.

Study of the relation between the distribution of human capital and the distribution of income has a long history in economics. The topic was the subject of a heated debate between Edgeworth and Pareto over how the distribution of ability would translate into the distribution of income. The early literature is

[15] Whenever δ is above the min-max majority, then there will be a set of unbeatable proposals. It is then important to understand the extent of the indeterminacy: how different are the various proposals which are undominated under $d(n, \rho)$-majority rule? A preliminary analysis of this issue is provided in Caplin and Nalebuff (1988). For $\rho \geq 0$ we may use these arguments to show that the set of undominated proposals is a "small" subset of the Pareto optimal set. All unbeatable proposals are close to the mean voter's most preferred outcome and the set shrinks as the dimension of the issue space is increased.

well-summarized in the survey article by Chipman (1976).[16] One reason for the amount of heat generated by this topic is the close relation to ethical questions. Some have argued that dispersion in income is largely the result of luck and prejudice. Others view the degree of inequality as a result of standard market forces and the heterogeneity in talents. To judge this issue clearly, it is important to study the connection between worker talents and the distribution of income.

Tinbergen (1956) and Roy (1950) pioneered the now-standard model of human capital and job selection. Each worker is characterized by an n-dimensional vector of skills. This bundle of skills is valued differently in each of the m different sectors of the economy. Specifically, we follow Mandlebrot (1962) and Heckman and Honoré (1990) in using a generalized Roy model of income determination. In our model, each sector has a different linear payment schedule for worker skills.[17] In addition, there may be a sector specific lump-sum payment (l_j) that is independent of the worker's type. The wage for a worker with skill bundle α in sector j is

$$\alpha \cdot w_j + l_j.$$

For example, an orchestra pays people primarily for their musical ability while a sports team places its emphasis on athletic ability when determining compensation. For those on welfare, wages may consist of only a lump sum payment, $w_j = 0$, $l_j > 0$.

Self-selection complicates the relationship between skills and income. Each individual chooses the job that maximizes income. As a result, instead of being linear, income is a convex function of worker skills. Define the maximal income for an α-type worker by $W(\alpha)$:

$$W(\alpha) = \max_j \alpha \cdot w_j + l_j.$$

PROPOSITION 4: *$W(\alpha)$ is convex over $\alpha \in R^n$.*

PROOF: Consider three workers, α_0, α_1, and their weighted average, α_λ. If all three work in the sector optimal for the α_λ type, then earnings will be a linear function of α. Convexity follows as the extreme types, α_0 and α_1, have earnings bounded below by these values (as either of them might earn an even higher income in another sector). *Q.E.D.*

This convexity property allows us to analyze the cumulative density for the economy-wide distribution of income. The set of workers who earn less than Y_λ contains the Minkowski average of the sets earning below Y_0 and Y_1. (As in the example in Section 2, the lower contour sets of a convex function have the

[16] We thank Andreu Mas-Colell for providing us with this reference.

[17] The original Roy model is restricted to two sectors and two worker attributes, with a joint lognormal distribution on these attributes. We allow for any number of sectors, any number of worker skills, and general ρ-concave densities. The lognormal distribution is handled as a special case and can be covered under the $\rho = 0$ class after a suitable change of variables.

required Minkowski averaging property.) We may then apply the Prékopa-Borell aggregation theorem to study the cumulative distribution of income, $F(Y)$,

$$F(Y) \equiv \int_{\{\alpha \in R^n:\, W(\alpha) \leqslant Y\}} f(\alpha)\, d\alpha.$$

PROPOSITION 5: *If the distribution of the n worker skills is ρ-concave, then the distribution of aggregate income, $F(Y)$, is $\rho/(1+n\rho)$-concave.*

This result holds for *any* set of wages in the different sectors. It is not possible to change this result by subsidizing the wages of jobs chosen by those earning the lowest wages.[18]

A central case is $\rho = 0$. Here, Proposition 5 proves that the distribution of income is log-concave. For example, this covers all cases in which skills are distributed according to any truncated normal distribution. While a normal distribution of skills leads to a truncated normal distribution of income within each sector, the aggregate distribution across all sectors is not normal, and deriving it in closed form is intractable. The preservation of log-concavity in the presence of self-selection is far from obvious.

Proposition 6 demonstrates another feature of the economy-wide distribution of income: there is a bound on the distance between mean and median incomes. The result is a direct implication of Proposition 3.

PROPOSITION 6: *If the distribution of worker skills is ρ-concave in R^n, then at the mean income $\bar{Y}$,*

$$F(\bar{Y}) \geqslant \left[\frac{1}{1+\rho^*}\right]^{1/\rho^*}, \qquad \rho^* = \frac{\rho}{1+n\rho}.$$

In particular, for $\rho \geqslant 0$ we can say that at least $1/e$ of the population must earn below average incomes.

Note, there is no upper bound on the fraction of the population that may earn less than average incomes as we are not able to conclude that $1 - F(Y)$ has any ρ-concavity properties. However, the intra-sector income density directly inherits ρ-concavity from the density of worker skills. Hence for sector i, both $F_i(Y)$ and $1 - F_i(Y)$ will be ρ^*-concave. By analogy to Proposition 6 we may conclude that within each sector at least $1/e$ and no more than $1 - 1/e$ fraction of the population earns below average incomes.

It is possible to extend Propositions 5 and 6 to cover densities over worker skills which are not log-concave and do not even satisfy A1. An important example is the multi-dimensional lognormal distribution. The linear characteristics model with a log-normal distribution of skills is equivalent to a transformed

[18] The income distribution results also characterize the shape of the supply function for unemployed workers as a function of the welfare payment. The log of the supply function is concave; this implies that the proportionate increase in voluntary unemployment is a decreasing function of the "unemployment" wage or welfare payment.

model where the wage for a worker of type α in sector j is

$$\sum_i e^{\alpha_i} w_{ij} + l_j,$$

where now α is normally distributed. Note that the maximal wage of an α type is still convex in α as the exponential transformation is a convex function. Hence, Proposition 5 continues to apply and the log of the cumulative distribution of income is still concave. What makes this example remarkable is that the density of income is very poorly behaved: it need not even be single peaked (see Heckman and Honoré (1990)) nor is the cumulative distribution within a sector necessarily log-concave.

Log-concavity of the cumulative distribution allows us to characterize various measures of income inequality. For example, Prékopa's theorem yields a simple proof of the result that with log-concavity, there is a rising gap between the wealthiest and the average as we move up the income distribution. This and related properties are also demonstrated in Heckman and Honoré (1990).

PROPOSITION 7: *The log-concavity of $F(Y)$ implies that the gap between income level Z and the conditional expectation of income given that $Y \leqslant Z$ is an increasing function of Z:*

$$\frac{d(Z - E[Y|Y \leqslant Z])}{dZ} \geqslant 0.$$

PROOF: Integration by parts show this gap equals

$$G(Z) = \int_0^Z F(\zeta)\, d\zeta / F(Z).$$

Since F is log-concave, we can apply the Prékopa theorem once more to claim that the integral of $F(\zeta)$ is also log-concave in Z:

$$F'(\zeta) \int_0^Z F(\zeta)\, d\zeta - F(\zeta)^2 \leqslant 0.$$

Differentiation of $G(Z)$ shows that the above inequality directly implies that $G'(Z) \geqslant 0$. *Q.E.D.*

Finally, there is an analogy between the number of firms competing for consumers and the number of firms competing for workers. In our companion paper, we show that with two-dimensional products, the average number of neighboring firms is six. Correspondingly, if wages are based on only two worker characteristics then the average industry competes with only six others for workers at the margin. In a one-dimensional model, the restriction is even stronger: each firm faces only two neighbors. This illustrates the importance of using an n-dimensional framework in which these restrictions are absent.

8. CONCLUSIONS

This paper establishes a multi-dimensional analog of Black's (1948) median voter result. We provide conditions under which the *mean* voter's most pre-

ferred outcome is unbeatable according to 64%-majority rule. The weaker restrictions supporting this result generalize Caplin and Nalebuff (1988).[19] Whenever the distribution of voter preferences is log-concave, we show that the preferences of the average voter must lie in a central position, viewed from any perspective. This limits the size of a coalition that favors change in any particular direction away from the mean voter's most preferred outcome. The idea that a distribution restriction places the mean in a central position is useful in other contexts. For example, we show how the mean voter theorem translates to a statement about the distribution of income around its mean.

To prove these results, this paper introduces to economics a mathematical aggregation theorem due to Prékopa and Borell. This new approach to aggregation has additional applications to such diverse topics as maximum likelihood estimation and search theory. Our companion paper, Caplin and Nalebuff (1991), shows how application of the Prékopa-Borell theorem provides the theoretical foundations for existence of equilibrium in imperfectly competitive markets.

Department of Economics, Columbia University, 116th and Amsterdam Ave., New York, NY 10027, U.S.A.

and

Yale School of Organization and Management, Box 1A, New Haven, CT 06520, U.S.A.

Manuscript received January, 1989; final revision received July, 1990.

REFERENCES

AVRIEL, M. (1972): "*r*-Convex Functions," *Mathematical Programming*, 2, 309–323.

BLACK, D. (1948a): "On the Rationale of Group Decision-Making," *Journal of Political Economy*, 56, 23–34.

——— (1948b): "The Decisions of a Committee using a Special Majority," *Econometrica*, 16, 245–261.

BONNESEN, T., AND W. FENCHEL (1987): "Theory of Convex Bodies," Moscow, ID: BCS Associates. Originally published as *Theorie Der Konvexen Körper*, 1934.

BRASCAMP, H. J., AND E. H. LIEB (1976): "On Extensions of the Brunn-Minkowski and Prékopa-Leindler Theorems, Including Inequalities for Log Concave Functions, and with an Application to the Diffusion Equation," *Journal of Functional Analysis*, 6, 366–389.

BORELL, C. (1975): "Convex Set Functions in *d*-Space," *Periodica Mathematica Hungarica*, 6, 111–136.

CAPLIN, A., AND B. NALEBUFF (1988): "On 64%-Majority Rule," *Econometrica*, 56, 787–815.

——— (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59, 25–59.

CHIPMAN, J. (1976): "The Paretian Heritage," in *Revue Européenne de Sciences Sociales et Cahiers Vilfredo Pareto*, ed. by G. Busino. Genéve: Droz.

CONDORCET, MARQUIS DE (1976): "Essay on the Application of Mathematics to the Theory of Decision Making," in *Condorcet, Selected Writings*, ed. by K. Baker. Indianapolis: Bobbs-Merrill.

[19] The applicability of 64%-majority rule covers the entire class of log-concave densities over voter preferences; this includes the truncated normal, exponential, and Weibull distributions. The result extends beyond the log-concave case and provides a bound on the min-max majority which depends on a concavity index for the distribution of preferences and the dimensionality of the preference space.

DAS GUPTA, S. (1980): "Brunn-Minkowski Inequality and Its Aftermath," *Journal of Multivariate Analysis*, 10, 296–318.

DHARMADHIKARI, S., AND K. JOAG-DEV (1988): *Unimodality, Convexity, and Applications*. San Diego, CA: Academic Press.

DONOHO, D. L., AND M. GASKO (1987): "Multivariate Generalizations of the Median and Trimmed Mean, I," Technical Report #133, Dept. of Statistics, U.C. Berkeley.

DRÈZE, J. (1974): "Investment Under Private Ownership: Optimality, Equilibrium, and Stability," in *Allocation Under Uncertainty: Equilibrium and Optimality*, ed. by J. Drèze. New York, NY: Wiley.

FLINN, C., AND J. HECKMAN (1983): "Are Unemployment and Out of the Labor Force Behaviorally Distinct Labor Force States?" *Journal of Labor Economics*, 1, 28–42.

GRANDMONT, J. M. (1978): "Intermediate Preferences and the Majority Rule," *Econometrica*, 46, 317–330.

GREENBERG, J. (1979): "Consistent Majority Rules over Compact Sets of Alternatives," *Econometrica*, 47, 627–636.

GRUNBAUM, B. (1960): "Partitions of Mass Distributions and of Convex Bodies by Hyperplanes," *Pacific Journal of Mathematics*, 10, 1257–1261.

HAMMER, P. C. (1960): "Volumes Cut from Convex Bodies by Planes," unpublished, made available by T. J. Smith, Dept. of Mathematics, University of Kalamazoo, Michigan.

HARDY, G. H., J. E. LITTLEWOOD, AND G. POLYA (1934): *Inequalities*. London/New York: Cambridge University Press.

HECKMAN, J. J., AND B. HONORÉ (1990): "The Empirical Content of the Roy Model" *Econometrica*, 58, 1121–1150.

HUBER, P. J. (1985): "Projection Pursuit," *Annals of Statistics*, 13, 435–475.

JEWITT, I. (1987): "Risk Aversion and the Choice Between Risky Prospects: The Preservation of Comparative Statics Results," *Review of Economic Studies*, 54, 73–85.

——— (1988): "Risk Aversion and the Static Portfolio Problem," University of Bristol Working Paper.

JOHNSON, N. L., AND S. KOTZ (1972): *Distributions in Statistics: Continuous Multivariate Distributions*. New York, NY: John Wiley & Sons.

KARLIN, S. (1968): *Total Positivity*. Stanford, CA: Stanford University Press.

KRAMER, G. H. (1973): "On a Class of Equilibrium Conditions for Majority Rule," *Econometrica*, 41, 285–297.

MANDELBROT, B. (1962): "Paretian Distributions and Income Maximization," *Quarterly Journal of Economics*, 76, 57–85.

MCFADDEN, D. (1981): "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data*, ed. by C. Manski and D. McFadden. Cambridge, MA: MIT Press.

MCKELVEY, R. D. (1979): "General Conditions for Global Intransitivities in Formal Voting Models," *Econometrica*, 47, 1085–1112.

NOLAN, D. (1989): "On the Min-Max Majority and Deepest Points," University of California, Statistics Dept. mimeo.

PLOTT, C. (1967): "A Notion of Equilibrium and its Possibility under Majority Rule," *American Economic Review*, 57, 787–806.

PRATT, J. W. (1981): "Concavity of the Log Likelihood Function," *Journal of the American Statistical Association*, 76, 103–106.

PRÉKOPA, A. (1971): "Logarithmic Concave Measures with Applications to Stochastic Programming," *Acta Sci. Math.* (Szeged), 32, 301–315.

——— (1973): "On Logarithmic Concave Measures and Functions," *Acta Sci. Math.* (Szeged), 34, 335–343.

RINOTT, Y. (1976): "On Convexity of Measures," *Ann. Probability*, 4, 1020–1026.

ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

RUBINSTEIN, A. (1979): "A Note About the 'Nowhere Denseness' of Societies Having an Equilibrium under Majority Rule," *Econometrica*, 47, 511–514.

SIMPSON, P. B. (1969): "On Defining Areas of Voter Choice," *Quarterly Journal of Economics*, 83, 478–490.

TINBERGEN, J. (1956): "On the Theory of Income Distribution," *Weltwirtschaftliches Archiv*, 77, 155–175.

TUKEY, J. W. (1977): *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.