Voting Systems, Honest Preferences and Pareto Optimality*

RICHARD ZECKHAUSER

Harvard University

"In a capitalist democracy there are essentially two methods by which social choices can be made: voting, typically used to make 'political' decisions, and the market mechanism, typically used to make 'economic' decisions."[1] To economists, the freely functioning competitive market mechanism has alluring structural properties.[2] If each individual follows his own self-interest in making his market decisions, an efficient outcome is achieved.

What of the possibility of constructing a voting mechanism that would achieve a collective political outcome that would have desirable properties that parallel those of the competitive market outcome? Can a voting system be impersonal and decentralized, allow each individual to act on his own behalf, and guarantee that a Pareto-optimal outcome will be achieved?[3]

* Kenneth Arrow, Robert Klitgaard, Pamela Memishian, Howard Raiffa, Thomas Schelling, Michael Spence, Milton Weinstein, and a referee for the *American Political Science Review* provided me with helpful comments. This research was supported by NSF grants GS-28626x and Gr-58.

[1] Kenneth Arrow, *Social Choice and Individual Values* (New York: John Wiley and Sons, Inc., 1964), p. 1.

[2] It is with some discomfort that I leave aside the problem of distribution in this paper. The issue is attacked directly in a later effort, "Risk Spreading and Distribution," Kennedy School of Government, Discussion Paper No. 10 (August, 1972). There I argue that a profound belief in the efficacy of the outcome of a perfect market gives one insights into distributional questions. One can consider the problem of drawing the social contract equivalent to that of a group of future citizens starting in some initial position of equal potential (they face the same lottery on future possibilities). They draw up a contingent claims agreement, the magnitudes of payment to depend on their future fortune in securing endowments of goods and capabilities (that is the states of the world). In retrospect, we may decide that it is our duty to redistribute in the manner that would be prescribed by the hypothetical agreement that antedated our present position. For a philosophically compatible, but perhaps policy inconsistent view, see John Rawls, *The Theory of Justice* (Cambridge, Mass., Harvard University Press, 1971).

[3] Thomas Schelling has commented at great length, interpreting and questioning the significance of Pareto optimality as a criterion for social choices. He states, "the significance of failure to achieve Pareto-optimality merely means that the situation is improvable . . . it is the degree of improvability that determines how important the nonoptimality is. Pareto optimality is not like virginity or justice. . . . Small departures are of small interest, large departures are of large interest." Personal communication, July 12, 1972.

Economists have long recognized that Pareto optimality by itself is a hopelessly incomplete guide to normative and descriptive considerations of social choice. An indication of this author's response is given by the conclusion to this paper.

A Desirable Voting Scheme

In the market, an individual indicates his preferences through his purchases and sales. In general, these tracings are insufficient to define an individual's complete preference mapping, but from the standpoint of efficiency this creates no difficulties. With merely the limited amounts of information fed to it, the market leads to a Pareto-optimal outcome. (It should be noted that no individual has an incentive to disguise his preferences.)

A voting system in theory could operate in the same decentralized fashion. Each individual would feed in some indication of his preferences, and the processing mechanism would employ this information to yield an outcome. To see whether such a procedure should be established as an ideal, we must determine how it can perform along a number of significant dimensions.

Major Characteristics of Social Decision Procedures

The flow diagram below provides the context for the major topics discussed in this paper. (At a later juncture, it will serve to summarize its principal results.) These topics are indexed A through F.

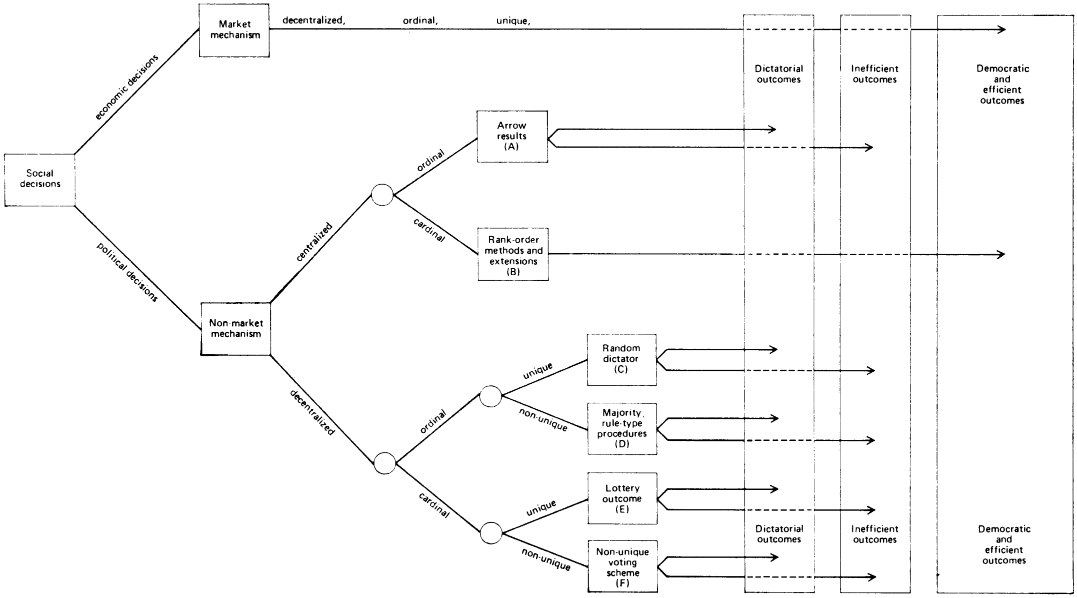
Social decision problems are first classed as economic or political. Economic decisions are handled by the market mechanism. Under the appropriate competitive conditions, the market outcome will meet some preassigned standards of acceptability. The principal standard to be met is efficiency. The measuring rod for this standard is Pareto optimality.

There are myriad situations in which the very restrictive competitive assumptions are not met. For these situations, much the subject of current economics debate, the unhindered play of the market will not lead to an attractive outcome. Perhaps the establishment of new areas of property rights, or innovative additions to present market mechanisms will set things right. Whatever, this is not material to the present discussion.

These situations would be of import for this paper if they were treated as political decisions. They then would be carried down the lower branch of the flow diagram, where they could serve as further exemplars for the analysis of that branch.

Political Decisions and the Individualist Approach

In evaluating outcomes in the economic arena,



Context and Summary Diagram

we rely on consumer sovereignty. Each individual is entitled to judge what is best for himself, and these individual judgments form the basis for assessments of social welfare. This paper carries over this individualist ethic to the evaluation of all social choices, including those that are conventionally considered to be of a political nature. Thus individuals' preferences form the basis for the social decisions considered here. In social decisions more than one individual is involved. Together, social decision and consumer sovereignty exclude from consideration any form of paternalistic rule.

The ground rules then are set. We are searching for a procedure that enables us to proceed from information on individuals' preferences to a social outcome that meets some conditions of acceptability. The ultimate objective is to discover a decentralized procedure, one which consciously or automatically takes the (perhaps incomplete) information provided by individuals' ballots and yields an acceptable social outcome.

Centralized Decision Procedures

For purposes of comparison, I shall begin with a look at the success or nonsuccess of centralized procedures. With centralized decision, a single decision authority with accurate knowledge of individuals' preferences selects an outcome.

Centralized Decision, Ordinal Preferences (A)

At first glance it would seem that centralized decision procedures hold out promise of great success. Yet Kenneth Arrow, in a much celebrated work, has demonstrated that there is no general procedure to pass from individuals' preferences to social orderings that meets certain standards of attractiveness.[4] The power of the Arrow result derives from the fact that the standards he imposed were both innocuous in appearance and intuitively appealing.

Arrow sets forth four standards. In unrigorous language, they are: (1) The procedure must include all logically possible combinations of individuals' orderings. (2) It must lead to Pareto-optimal outcomes. (3) The choice between any two alternatives cannot be influenced by the presence or nonpresence of a third alternative. (4) No individual can always secure his choice regardless of the preferences of others. A later expositor has provided a mnemonic for these properties:[5] unrestricted domain, U; Pareto principle, P; independence of irrelevant alternatives, I; nondictatorship, D. Dealing with the simplest case of interest (at least two individuals and at least three alternatives), the famous Arrow result is that there is no social welfare function that satisfies U, P, I, and D.

Let us stop for a moment to get our bearings in the flow chart. We should like all of our social decisions to apply over an unrestricted domain. Condition U is imposed at all nodes in the flow diagram and throughout this paper. Next observe

[4] For the conditions, see K. Arrow, *Social Choice*, pp. 22–33.

[5] Amartya K. Sen, *Collective Choice and Social Welfare* (Holden-Day, Inc., San Francisco), pp. 41–42.

that the "Arrow result" box lies on the ordinal branch. The independence of irrelevant alternatives condition, I, requires that the centralized decision procedure rely on no more than individuals' ordinal preferences. The impossibility of simultaneously satisfying all four Arrow conditions is indicated by the paths of the arrows emanating from the box. They are stopped either by inefficiency (nonsatisfaction of P) or dictatorship (nonsatisfaction of D).

Centralized Decision, Cardinal Preferences (B)

Somewhat different formulations of the centralized social choice problem will lead to success stories. A number of analysts, some of whom substantially predate Arrow, have proposed that acceptable centralized decision procedures can relax one or another of the conditions he imposed. Perhaps the most frequently cited suggestion is that the independence of irrelevant alternatives condition, I, be softened or eliminated. The objective underlying this relaxation is to secure some cardinal measurement on individuals' intensities of preferences.

By way of example, more than two centuries ago, Borda proposed the rank-order method of election, a method now frequently employed to establish national rankings of college athletic teams. Under this system, preassigned weights are given to voters' first, second, third choices, etc. Sum each candidate's scores over all voters: highest aggregate score wins. The procedure is neither dictatorial nor inefficient.[6] Reference to the flow diagram shows that its emanating arrow penetrates to the rightmost rectangle.

Are rank-order systems ethically appealing? Neither economic theory nor any other axiomatic system can give us any insight into this question. Each analyst of social choice must decide for himself whether in making the choice between A and B, it is of concern how each fares with the voters not only with respect to the other, but also with respect to C.

We need dwell no further on the appropriate constraints for centralized procedures; decentralized procedures represent the principal focus of this paper. Thus far we have been merely setting the stage and filling in the complementary boxes.

Decentralized Decision Procedures

The remainder of this paper will follow out the implications of and possible standards of performance for a variety of decentralized decision procedures. It will index them, as it did the centralized procedures, by the capital letters in the flow diagram.

The analysis that follows is largely nonmathematical and thus almost by necessity, unrigorous. Other theorists, no doubt, will be able to reconstitute these results in more formal, perhaps stronger contexts. The purpose of this paper is to sketch some new methods and results, and the framework in which they apply.

With centralized decision, each voter's preferences are assumed to be known by the social choice mechanism. With decentralized decision, knowledge of preferences is gained only indirectly. Under decentralized procedures each voter fills out a ballot; he is assumed to follow as best he can his own self-interest. These ballots are processed by some mechanism to produce a social choice. Depending upon the particular choice mechanism that is in operation, and the situation in question, a voter's ballot may or may not be a complete and truthful indicator of his preferences. This observation provides a hint about where we are going. A whole new class of problems arises when we move from centralized to decentralized decision procedures. Essentially they revolve around getting individuals to reveal their preferences. The unfortunate result, as we shall see later, is that the very structure of social choice schemes that elicit true preferences on ballots leads to inefficiencies in outcomes.

An abstract representation of a decentralized choice procedure will facilitate discussion. There are

m voting individuals indexed over i, and
n alternatives indexed over j.

An individual, i, fills out x_i as the ballot indicating his preferences. The voting system processes the received ballots, $X=(x_1, \cdots, x_i, \cdots, x_m)$ to arrive at a social choice. A social choice is represented by the probability vector $P(X)=(p_1(X), \cdots, p_j(X), \cdots, p_n(X))$ whose elements indicate the probability that an alternative is selected. In the language of mathematics, a voting mechanism maps X into P.

Most of the familiar voting procedures are deterministic. Given X, a single alternative is selected with certainty; one element of P is unity, the rest zero.

Later sections of this paper discuss voting pro-

[6] Following Pareto optimality, efficiency requires that a switch away from the selected candidate would lower the welfare of at least one voting individual. If the selected candidate had the highest total score, then there could be no other candidate who received an equally high individual score from each voter. A switch from the man with the highest total would therefore involve a loss for at least one voter.

Special difficulties arise if lotteries on alternatives are permitted; and rank-order schemes need not lead to Pareto-optimal outcomes. See Richard Zeckhauser, "Majority Rule with Lotteries on Alternatives," *The Quarterly Journal of Economics*, 83 (November, 1969), 696–703.

cedures that may lead to randomized outcomes. These outcomes are analyzed and evaluated before the final results are known, before the lottery is conducted. Randomized outcomes may lead to gains in efficiency. More important for the purposes of this paper, they may help insure that individuals mark their ballots with their true preferences.

To evaluate probabilistic outcomes, a voter must be able to evaluate outcomes in a nondeterministic world. Even if the world is deterministic, there may not be full information, in which case the voter may be making a decision under uncertainty. Fortunately, there is a well-developed body of literature on prescriptive decision making under uncertainty. We shall follow the major track of that literature.[7]

A rational decision maker under uncertainty is assumed to assign a utility value to each possible outcome. The paradigm decision problem requires that he select one among alternative lotteries where each lottery is described by the probability vector it offers for the different outcomes. In choosing among lotteries, the rational decision maker selects the one that offers him the highest expected utility value.

Notice that judgments of welfare are made here on an ex ante basis, before lotteries are resolved. Thus certain social choices will be ruled non-Pareto optimal, even though they will eventually lead to situations that considered in isolation are Pareto optimal. An example displays the properties of this before-the-fact evaluation procedure. Two risk-averse individuals each have $10,000 in the bank. They are given the opportunity to gamble their total savings on the flip of a coin. Both would decline as they both prefer their present situation to participation in the lottery. If they were forced to participate in the lottery, it would be a non-Pareto-optimal situation. This despite the fact that if the lottery were conducted and one man ended up with $20,000 and the other with nothing the outcome would be Pareto-optimal.[8]

[7] For a good introduction to this subject, see Howard Raiffa, *Decision Analysis* (Reading, Mass.: Addison-Wesley, 1968).

[8] The ex-ante expected utility approach can be employed to make evaluations of social welfare. See Richard Zeckhauser, "Determining the Qualities of a Public Good—A Paradigm on Town Park Location," *Western Economic Journal,* 11 (March, 1973), 39–60. For an application of this principle in the social decision context, see Richard Zeckhauser, "Majority Rule With Lotteries on Alternatives." Peter Fishburn formalizes this argument in "Lotteries and Social Choices," *Journal of Economic Theory,* 5 (October, 1972), 189–207. Peter C. Ordeshook employs the principle to define candidates' mixed strategies as lotteries on social states should they get elected. See "Pareto Optimality in Electoral Competition," *The American Political Science Review,* 65 (December, 1971), 1141–1145.

To conduct an analysis of situations of this sort requires that individuals define what are called von Neumann-Morgenstern utilities for each possible outcome.[9] The utility values for individual i are indicated by the v_j^i's; they are normalized with a value of 1 assigned to the most preferred outcome, and 0 to the least.

The voting mechanism, however, may seek to elicit something much less than the individual's full cardinal preference structure. Majority rule procedures, for example, merely ask the individual which is his most preferred alternative.

It is useful to classify decentralized decision procedures on a couple of dimensions. Decentralized procedures, like those that are centralized, can rely on cardinal or merely ordinal indications of individual's preferences. The two alternatives are indicated in the first branching in the flow diagram.

Unique Voting Schemes

The second branching reflects a consideration that does not arise with centralized procedures; for with them, individuals' preferences are assumed to be known. With decentralized procedures, however, each individual ballots in his own self-interest. This raises the possibility that with some of these procedures and some situations an individual working in his own behalf will find it desirable to disguise his preferences. This may be unfortunate. It may be important for the apparatus that is processing the votes to be able to determine the exact preferences of an individual from his markings on the ballot.

If an individual will mark his ballot the same way for two or more preference structures, then the ballot is not unique.

Definition: A voting system is called *unique* if a voting individual will mark his ballot in a different manner for every possible configuration of his preferences.

Decentralized Decision, Ordinal Preferences, Unique (C)

If merely ordinal preferences are required, it would seem the task of constructing a unique scheme would be somewhat simpler. In many real world situations things are simpler still. Individuals are asked to indicate only a small portion of their ordinal preference structure, their first-place choice. This is the customary procedure in the United States when we are selecting a single individual for an office: the presidency, a mayorship, the local dog catcher.

[9] Appropriate historical attribution would give much of the credit for this utility concept to Frank P. Ramsey.

A traditional majority-rule decision procedure will lead to honest revelation of preferences if there are but two alternatives. But when the competitors number three or more such honesty can no longer be assured. For example, if in a three-man race your preferred candidate A is likely to receive little support, you may choose to vote for your second choice rather than waste your vote on a lost cause. If your ordering is A-B-C, you will vote for B. You would vote the same way if your ordering were B-A-C or B-C-A. The voting procedure is not unique, but this should have been expected. There are six possible orderings and only three possible ballot markings. There is no way to have a unique marking for each ordering. What is disturbing is that the system is not unique with respect to the first-place preference, and it is information on that preference that the system was attempting to solicit. Considerations such as these lead us to a modified concept of uniqueness, one that applies in a more restricted context.

Definition: A voting system is called *unique with respect to a property of a preference ordering* if a voting individual will mark his ballot in a different manner for every possible configuration of that property.

For example, a voting system is unique with respect to first-place preferences if a voting individual will mark his ballot differently depending upon which candidate he likes best.[10] Majority rule does not qualify.

Does a voting system exist that is unique with respect to first-place preferences? Yes, there is a type of system that meets this requirement. It is most easily understood with the aid of a mechanical example that has the required intrinsic properties.

The Random Dictator System: A Voting System That is Unique with Respect to First-Place Preferences

Each individual writes the name of a candidate on a ballot. The voters' ballots are collected and placed in a revolving drum. After shuffling, a ballot is chosen at random. The name on the chosen ballot is the elected candidate. In effect, the voter whose ballot is chosen dictates the outcome, hence the name of the system. (Matters would be just the same if a voter were first selected at random, and then the selected voter were asked to choose the elected candidate.)

The random dictator system is not deterministic. If an alternative j is marked as first choice on k ballots, p_j will equal k/m. That is, a candidate will have a probability of election that is proportional to the number of first place votes he receives. The p_j's will take on fractional values (except in the special case where there is a unanimous first choice); the system will not be deterministic.

We shall see here, and again later with cardinal preferences, that the consideration of probabilistic outcomes is essential if a voting scheme is to be unique. Consider the voter's behavior in the random dictator model. He assigns utility values v_A, v_B, and v_C to the three candidates. Why should he vote for his favorite, A? Assume that instead he had voted for B. This would reduce the probability that A gets elected by $1/m$; it would increase the probability that B got elected by $1/m$. The net change in the expected utility to the voter would be $(1/m)(v_B - v_A)$. By assumption, $v_B < v_A$; the voter would lose by switching; he also would lose if he decided to choose C. The key element here is that when the individual switches his vote the probability gain for the switched-up candidate is just equal to the probability loss for the switched-down candidate. This implies that the probability of election of the third candidate remains constant when a vote is switched between the other two. This absence of third-candidate effect is essential. Otherwise, the decision on whether to vote for A or B might depend on how the vote affects the probability that C gets elected.[11] A formal generalization of this argument requires a new concept.

Definition: A voting system will be said to be *probabilistically linear* if it is of the following form. Let S be the matrix of votes, where element $s_{ij}=1$ if voter i submitted his ballot for j, and 0 otherwise. The vector P that gives the probability that each candidate is elected is calculated $P = wS + e$. Here e is a row vector giving exogenous contributions to the election probabilities, and w is a row vector of non-negative weights summing to $1 - \sum_j e_j$.

Theorem I: If a decision procedure that solicits individuals' first place preferences is to be unique, it must be probabilistically linear.

Note that probabilistic linearity is equivalent to the following property: For any voter i and for any pair of candidates D and E, switching his vote from D to E must increase the probability that candidate E gets elected by the amount w_i, de-

[10] If we extend this concept to the market setting, we discover that market procedures are unique with respect to the information required for efficiency: individuals' marginal rates of substitution and producers' marginal rates of transformation among valued goods.

[11] Say v_A is only slightly greater than v_B, but much greater than v_C. If C had less probability of being elected if the individual voted for B rather than A, it might be reasonable for him to vote contrary to his true preferences.

crease the probability that D gets elected by the amount w_i, and leave all other election probabilities unchanged.

The proof of the theorem can be demonstrated with a three-man example. For notation, let $\Delta^{F}_{D \text{ to } E}$ represent the change in the probability that F is elected if voter i switches his vote from D to E. Consider the previous example where the present vote is for A. Necessary conditions for probabilistic linearity are that

$$\Delta^{A}_{A \text{ to } B} = - \Delta^{B}_{A \text{ to } B} = \Delta^{A}_{A \text{ to } C} = - \Delta^{C}_{A \text{ to } C}$$

and that

$$\Delta^{C}_{A \text{ to } B} = \Delta^{B}_{A \text{ to } C} = 0.$$

Given that the sum of the mutually exclusive probabilities of election must add to 1 as they did when the vote went to A, there are the conservation equations

$$\Delta^{A}_{A \text{ to } B} + \Delta^{B}_{A \text{ to } B} + \Delta^{C}_{A \text{ to } B} = 0$$

and

$$\Delta^{A}_{A \text{ to } C} + \Delta^{B}_{A \text{ to } C} + \Delta^{C}_{A \text{ to } C} = 0.$$

To see that $\Delta^{C}_{A \text{ to } B}$ must equal 0, assume it were negative. The voting system must handle any set of preferences. Thus, let the voter strongly dislike C, say $v_A - v_C = k$, where k is a large number, and let him slightly prefer A to B, $v_A - v_B = \epsilon$. Despite this preference, the voter will switch his ballot to B, as his net utility gain will be $-\Delta^{C}_{A \text{ to } B}(k-\epsilon)$, a positive number. Keep the same ordinal preferences, but let the gap between v_A and v_B be k, with the gap between B and C being the minimal ϵ. Then the individual will choose A over B, the utility gain once again being $-\Delta^{C}_{A \text{ to } B}(k-\epsilon)$. Thus, unless $\Delta^{C}_{A \text{ to } B}=0$, the voter may ballot in two different ways, despite the fact that his ordinal preferences are the same in both instances. This contradicts the assumption that the voting system is unique. Therefore $\Delta^{C}_{A \text{ to } B}=0$ as claimed.

An identical argument shows that $\Delta^{B}_{A \text{ to } C}=0$. These results together with the conservation equations show that $\Delta^{A}_{A \text{ to } B} = -\Delta^{B}_{A \text{ to } B}$ and $\Delta^{A}_{A \text{ to } C} = -\Delta^{C}_{A \text{ to } C}$. To show that $\Delta^{A}_{A \text{ to } B}$ must equal $\Delta^{A}_{A \text{ to } C}$, merely apply the argument above to show that whether you ballot for B or C cannot influence the probability of election for A if a voting system is to be unique with respect to ordinal preferences. The conditions of the theorem are thus demonstrated.

This theorem and the random dictator system can be extended to deal with further places in voters' ordinal rankings. Provide each voter q ballots for his first choice, r for his second, s for his third, etc., with $q>r>s$. The selection procedure is random as before.

Thus we find that only variants of the random dictator system will elicit ballots unique with respect to individuals' (portions of individuals') ordinal preferences. How well does the random dictator system perform? An example tells the sad news.

With the values shown in the box, Voter 1 will mark his ballot with A and Voter 2 will inscribe a C.

Utility Valuations in Two Voter World

	v_A	v_B	v_C
Voter 1	1	.90	0
Voter 2	0	.90	1

The resulting social choice when the random dictator system is employed will have the probability vector $P=(.5, 0, .5)$. The expected utility values for each of the voters will be .50. But this outcome is not Pareto optimal. Both voters would have preferred the social choice $P=(0, 1, 0)$.

The difficulty is apparent. The random dictator system does not take into account intensities of preference. (Indeed, no information is gathered on second and further preferences, but even if it were, this would not cover the intensities problem. With the same ordinal patterns, both voters could have assigned a value of .10 to v_B. In that case the social choice of the random dictator device would be Pareto-optimal, as would be any lottery on A and C.)[12]

To sum up, it is interesting to observe that there exists a class of unique ordinal voting schemes.[13]

[12] For further discussion of this matter, see R. Zeckhauser, "Majority Rule with Lotteries on Alternatives."

[13] This paper is addressed to situations where a single outcome is selected, and all individuals vote to help determine that outcome. The model does not apply, for example, to the selection of a legislature. (With a stretch of the imagination, we might construe voters to be choosing among all possible legislatures, but usually an individual voter ballots to help determine only a small part of that body.) Strict proportional representation is the closest (but still imperfect) legislative equivalent to the random dictator system. The number of seats secured is proportional to the number of votes received.

Proportional representation will guarantee honest balloting if two conditions are satisfied. (1) Each voter's utility for the legislature must be an additive separable function of his utilities for each elected representative. (Otherwise, his preferred vote might depend on the candidates others were voting for.) (2) All ballots cast in the first round must count in the determination of the final legislature. This requires that representation be broken down into arbitrarily small units, and that there be no elimination of candidates. (Condition (2) can be violated if there are strict party interests, and if each voter is indifferent among the candidates of his party.) See Mark Thompson and Richard Zeckhauser, "Proportional Representation," mimeo (1971). A referee suggests that Kramer also has unpublished work on this subject.

We now know that such schemes perform terribly inefficiently. Leave aside the adverb, and this lesson could have been derived directly from the Arrow results.

Decentralized Decision, Ordinal Preferences, Nonunique (D)

This next class of social choice procedures merits discussion primarily because it is the one most frequently in use in the United States. Majority and plurality rule both fall under this category. The Arrow results tell us that these schemes cannot meet his four standards either.

One difficulty with these procedures, as is well known, is that they lead to intransitive choice patterns for certain combinations of individuals' preference orderings. Transitivity is required if we are to have a social ordering. The only way around the problem is to rule out patterns of individuals' preferences that do not satisfy the single-peakedness condition.[14] To exclude in this way would be to violate Arrow's condition *U*.

Now that we are working on a decentralized basis, further problems may arise. The vote processing mechanism may not even be able to divine individuals' true preference orderings. It turns out, fortuitously, that single-peakedness is also sufficient to eliminate any incentive for a voter to disguise his preferences. Getting rid of one problem automatically eliminates the other. Whenever the structure of voters' preferences is such that majority rule can be guaranteed to work on a centralized basis, it can be guaranteed to work in a decentralized format as well.

Standards for Decentralized Schemes

One lesson we have learned thus far is that ordinal preferences do not provide sufficient information to enable social choices to meet Arrow's standards. For centralized procedures, we saw that information on cardinal preferences is sufficient to guarantee an efficient, nondictatorial outcome. The key question for the remainder of this paper is whether we can find decentralized procedures that perform satisfactorily when the balloting system elicits cardinal indications of the voters' preferences.

Satisfactoriness, like beauty, is in the eyes of the beholder. My criteria, consistent with the earlier analysis, are Pareto optimality and nondictatorship. The first concept is unambiguous; the latter is subject to a variety of interpretations. I shall provide one definition of dictatorial outcomes that enables me to achieve my other results. The reader may come up with other definitions that serve equally well.

Dictatorial Outcomes—What They Are and Why They Should Bother Us

Most discussions of social choice are concerned with the selection of a candidate in a single election, or an outcome in a particular set of circumstances. Given the restricted domain of such a social decision, it might be perfectly acceptable to reach an outcome that catered heavily to one individual's or group's preferences and little or not at all to those of others. (Surely, this provides the normative justification for such practices as logrolling.)

Other social decisions can be of much greater consequence. Indeed at the most general level, we can conceive of a grand social decision as selecting a single point from the entire space of social states. It is with reference to momentous social decisions of this sort that we impose distributional requirements on the outcomes of social choice processes. We should like to observe outcomes that demonstrate that the preferences of more than one individual have been taken into account.

Even if a series of social choices is made sequentially or in a variety of separate arenas, we might conceive of the combination of the voters' expressed preferences on all issues and the social choice mechanism as selecting one point in the space of social states. Thus even if a great number of social choices is to be made on a one by one basis, we should like to see the total social outcome reflect the fact that all individuals have participated in the selection process.

One indication that a group had been excluded from positive consideration in the social choice procedure would be that it selected an outcome that was most unattractive for that group. A more precise understanding of what is meant by unattractive is given by the term Pareto-terrible.

> Definition: An outcome is called *Pareto-terrible* for a group of individuals if it is a strictly constrained minimum: if there exists an outcome that is preferred by all members of the group, and if there is no outcome that is worse for some member without being better for some other member.

An outcome is Pareto-terrible for a one-man group, an individual, if it is his least preferred outcome. With lotteries permitted there will always be outcomes that are not Pareto-terrible for any reference group.

I shall employ the Pareto-terrible concept as a cornerstone of my definition of dictatorial outcomes.

[14] See A. Sen, *Collective Choice*, pp. 166–168 for a discussion of single-peakedness where individuals' preferences are defined over a single dimension. There is no equivalent to the single-peakedness criterion if two or more dimensions enter individuals' preferences; cyclical majorities will be the rule.

Definition: A voting scheme is *dictatorial* for a set of voting individuals preferences if it selects an outcome that is Pareto-optimal for one nonempty subset of individuals and is Pareto-terrible for the complementary nonempty subset of individuals.[15]

With this definition in tow, we can proceed on our search. Exposition is simplest if we step slightly out of order to the bottom of the Context and Summary Diagram and address first the question of honest balloting.

Decentralized Decision, Cardinal, Nonunique (F)

Nondictatorial, Pareto-optimal social choices cannot be made unless, in general, the voting scheme is unique with respect to individuals' cardinal preferences. In particular instances nonuniqueness may not matter. For every nonunique ballot marking for individual 1, however, there will be a set of preferences for individual 2 such that Pareto optimality cannot be assured.

A three-candidate example will reveal the difficulties. Assume that individual 1 marks his ballot x' whether his preferences are (1, .60, 0) or (1, .70, 0) for candidates A, B, and C respectively. Individual 2 has preferences (0, .35, 1). If individual 1's FIRST set of preferences hold, candidate B will have zero probability for all Pareto-optimal choices. With individual 1's SECOND preferences, all Pareto-optimal probability vectors have either A or C receiving zero probability.

Pareto Optimality of Different Probability Vectors

Configuration of Individual 1's Preferences	Probability Vectors ($A+$ indicates positive probability)						
	1, 0, 0	0, 1, 0	0, 0, 1	+, +, +	+, +, 0	+, 0, +	0, +, +
FIRST	yes	no	yes	no	no	yes	no
SECOND	yes	yes	yes	no	yes	no	yes

Except for the first and third listed vectors, no probability vectors are Pareto-optimal for both of individual 1's possible preference configurations. The problem is simple. The mechanism that processes the votes is unable to distinguish between two situations which require different outcomes to be Pareto-optimal.

The 1, 0, 0 and 0, 0, 1 vectors are ruled out because they are dictatorial. They allow one voter to get his first choice at the expense of the other who receives the outcome that for him is worst. This would be particularly objectionable if, as is quite possible, the voting scheme were selecting a single point from the space of social states, i.e., if all social decisions were being simultaneously resolved.[16]

To achieve Pareto optimality, any voting scheme that is nondictatorial for this set of preferences must be able to distinguish between FIRST and SECOND configurations of individual 1's preferences. An equivalent "counterexample" demonstration can be arranged for any indistinguishable pair of individual 1's preferences. The three-alternative case covers all essential complexities.

First consider the case where individual 1's indistinguishable configurations have different alternatives as the most preferred. Let individual 2 be indifferent between the two sometimes-most-preferred-by-1 alternatives, with the third alternative less attractive. Which of two alternatives is Pareto-optimal depends on which of individual 1's preference configurations applies; there is no way to tell.

Matters are a bit more complicated if individual 1's indistinguishable orderings rearrange his second and third candidates or merely change their relative scores. To construct the counterexample first normalize the two individuals' utilities to 1 and 0. Add together 1's and 2's utilities for each of the three candidates. The candidate with the lowest total must receive zero probability in any nondictatorial Pareto-optimal outcome. (If the

[15] The concluding theorem of the paper employs the dictatorship concept negatively. It states that no system can simultaneously be nondictatorial and at the same time guarantee that other desirable standards are met. The more restricted the definition of dictatorship, then, the stronger my final result. This result would still be achievable if the term "Pareto-optimal" in the above definition were replaced with "unanimously most preferred." The argument for the definition in the text is that it captures more of what we may mean by dictatorship in a general context.

It should be noted that Pareto-terrible is not a perfect counterpart to Pareto-optimal. There are situations where all outcomes are Pareto-optimal and none Pareto-terrible. The converse is not true.

[16] On a less grandiose level, there are a number of ways that single social-choice decisions can be formulated to deal with a large number of issues. Issues can be compounded. The issues at hand might be whether Smith or Jones should be mayor and whether or not the new high school should be built. This can be viewed as a single decision with four possible outcomes. Alternatively, we might amalgamate issues by employing a lottery. A coin will be flipped to see whether a voting mechanism will select a mayor or decide on a high school. Here too there are four possible outcomes, each composed of two contingent decisions.

two voters have opposite orderings, then the extreme candidates may tie for lowest total. Then, at least one of these candidates will be excluded from any Pareto-optimal outcome.) Given the two different utility vectors for 1, select a utility vector for 2 so that (a) his first place choice is not the same as that of 1, and (b) the candidate receiving the lowest total will differ depending on which of 1's vectors applies. Such a selection will always be possible.[17]

It is true, of course, that for certain patterns of preference for 2, particular instances of nonuniqueness for 1 will not matter. This raises the question whether a voting individual can monitor other individuals' preferences or ballot markings to help select his own ballot marking. Such monitoring raises no problems so long as each individual's markings are unique, conditional on the ballot markings of others. For with this information, all individuals' accurate cardinal preferences can be deduced.

A definitional problem arises with the concept of uniqueness if voters monitor each others' ballots and only employ the same marking for two different preference orderings if they discover that it will not affect the set of Pareto-optimal outcomes. Notice that this monitoring would have to be most accurate, for as long as there is any positive probability that the other individual(s) have a "counterexample" set of preferences, Pareto optimality cannot be assured. In discussing uniqueness for the remainder of this paper, I shall assume either (a) that the required accurate monitoring is not possible, or (b) if it is possible, an individual who has an incentive to mark his ballot nonuniquely will sometimes do so in cases where other individuals have counterexample-relevant sets of preferences.[18]

This slightly tortuous argument leads to a general result that can serve as a building block for later arguments.

> Theorem II: For nondictatorial voting schemes, if the ballot processing mechanism is always to lead to Pareto-optimal outcomes, the voting individuals' ballots must be unique with respect to their cardinal preferences.

[17] Another numerical example may be helpful. Assume that 1 marks his ballot the same way whether his cardinal utility values are 1,.5,0 or 1,0,.2. Then if individual 2 has preferences 0,1,.9, for example, a nondictatorial Pareto-optimal outcome cannot be assured. The third candidate can never be included for individual 1's first set of preferences and must always be included for his second set of preferences.

[18] I strongly believe that it is possible to substantiate the arguments in the remainder of this paper without these assumptions. But the whole area is most elusive, and I have not been able to discover a convincing proof for this assertion.

Decentralized Decision, Cardinal, Unique (E)

To continue the thrust of our major investigation, we should like to know the properties of voting schemes that are unique with respect to cardinal preferences. In essence, we are asking: What voting schemes will get individuals to mark their true cardinal preferences? Unique schemes require one-to-one correspondences between ballot markings and preferences. If the mapping process is understood, we can say that in effect individuals are marking their true preferences.[19]

The required structural characteristics can be easily delineated in our three-candidate example. Before marking his ballot, the individual voter will consider his possible effects on the probability outcome vector. Matters are simplest if we start with the case of one individual.

What are the structural characteristics of a decision process that gets a single individual to reveal his true cardinal preferences? The individual is assumed to understand how the scheme works. The scheme is an announced function, s, that transforms the individual's ballot into a probability vector outcome. This process can be represented as

$$P = s(x).$$

The individual, knowing the form of s, will select x to maximize his expected utility. Representing his utility vector as V, and his optimal ballot as x^*, this can be expressed

$$\max_x s(x)V^T \qquad \text{is achieved at } x^*.$$

We wish to discover the properties of the function s that will lead

$$x^* \equiv V.$$

That is, we wish to construct a function s such that it is in the individual's best interest to vote his true cardinal preference.

First, the function must obviously assign probabilities in the same rank order as the utilities that are provided through the x vector. (For simplicity,

[19] It may be possible to construct voting schemes that are not unique for each individual taken alone, but are unique for all individuals taken together. This would require that individuals have some capability to monitor each others' preferences and ballot markings. For example, individual 1 might mark his ballot the same way when he has preferences V_1' and individual 2 marks X_2' as he does when he has preferences V_1'' and individual 2 marks X_2''. The processing system can take 2's markings into account in deciphering individual 1's preferences.

If each individual's cardinal preferences are to be revealed, it is sufficient that *given other individuals' markings* there be a one-to-one correspondence between an individual's markings and his preferences.

we consider the case without ties.) Otherwise, the individual would have an incentive to misreport which alternative he liked best and which least.

With no loss of generality, we can normalize and assign a utility value of 1 to the most preferred alternative and 0 to the least. Assign the value y to the intermediate alternative; thus $V=(1, y, 0)$.

The key question then is, what additional properties must $s(x)$ have so that

$$\max_z s(1, z, 0)\begin{pmatrix}1\\ y\\ 0\end{pmatrix} \quad \text{is achieved at } z = y$$

for all y, $0<y<1$.

The sum of the probabilities of the outcomes must equal 1. Thus, the function $s(x)$ is fully described by detailing the effect of the choice of z on p_1 and p_2. Represent these relationships

$$p_1 = f(z), \quad \text{and}$$
$$p_2 = g(z).$$

The individual's expected utility is

$$1p_1 + yp_2 + 0(1 - p_1 - p_2).$$

He will select z to maximize

$$f(z) + yg(z);$$

the third term vanishes because of the zero multiplicand.

I will deal with the situation in which the f and g functions are continuous and differentiable; my results can easily be extended to noncontinuous cases. First-order maximization procedures applied to the preceding expression reveal that

$$f'(z) = -yg'(z).$$

If it is to be in the individual's self-interest to select $z=y$, a significant property of f and g is dictated.[20] For all z

$$f'/g' = -z. \qquad [A]$$

Additional required properties are that for all z,

$$f''(z) + zg''(z) < 0,$$
$$g(z) < f(z),$$

and

$$g(z) > 1 - f(z) - g(z). \qquad [B]$$

[20] The uniqueness requirement is equally well satisfied so long as there is a one-to-one mapping between z and y. Any scheme that induces such a mapping has the key structural aspects that are discussed below in relation to this scheme.

The first requirement tumbles out of the second-order conditions that guarantee that the individual is at a maximum, not a minimum. The second and third requirements insure that the probability assigned to the highest value outcome is greatest, the second is second, and the lowest is lowest.

We should also like to have

$$f(1) = g(1),$$
$$g(0) = 1 - f(0) - g(0).$$

If alternatives are ranked the same, they should receive the same probability assessment.[21]

There is an infinite number of schemes that meet these requirements. One satisfactory scheme has the functions

$$p_1 = f(z) = .6 - .1z^2/2,$$
$$p_2 = g(z) = .25 + .1z,$$

and

$$p_3 = 1 - p_1 - p_2 = .15 + .1(z^2/2 - z).$$

Here $g'(z)=.1$; and to satisfy requirements

$$f'(z) = -zg'(z) = -.1z$$

for $z=y$. For all values of z, there must be the ordering $p_1>p_2>p_3$. This three-alternative scheme is illustrated in the figure below.

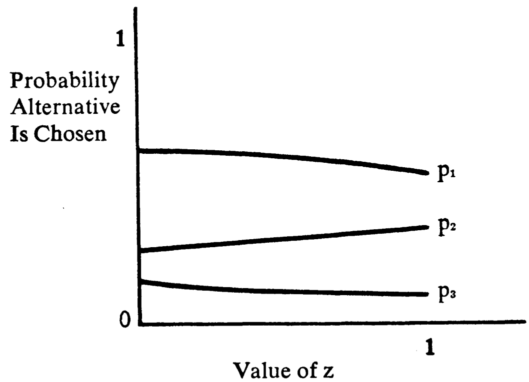


Figure 1. A scheme that leads an individual to express his true cardinal preferences.

	Value of z					
	0	.2	.4	.6	.8	1
p_1	.6	.598	.592	.582	.568	.55
p_2	.25	.27	.29	.31	.33	.35
p_3	.15	.132	.118	.107	.102	.10

[21] I am indebted to Kenneth Arrow for this point among others.

The procedure just employed to construct a unique voting scheme for cardinal preferences can be readily extended to the many-alternative case. The computations get somewhat messy as the number of alternatives increases, but no new conceptual problems arise.[22]

Pareto Optimality of the Unique with Respect to Cardinal Preferences Scheme

The scheme presented is unique with respect to the individual's cardinal preferences. Unfortunately, it is not Pareto-optimal. Pareto optimality would require that $p_1=1$. This misfortune, alas, is representative. With rare exceptions, schemes that are unique with respect to information required to make an efficient decision are by their very construction not Pareto-optimal.

The general difficulties that are encountered can be understood in the context of the simplest case where the difficulties matter, the case of two voters and three alternatives. (The one-voter case is not disturbing; Theorem II applies only to nondictatorial voting procedures. With one individual, dictatorship is acceptable; the cardinal preferences required by Theorem II are not needed to achieve a desirable outcome. Merely ask the individual which alternative he prefers and give it to him.)

Voting Schemes with Two Voters and Three Alternatives

With two or more individuals, as Theorem II tells us, efficient voting procedures that are democratic require cardinal preferences as inputs. The procedure must be unique with respect to individuals' cardinal preferences.

The insights we have derived from consideration of the one individual situation are readily extended. Look at the matter from the standpoint of voter 1. Assume that the vector giving voter 2's true cardinal preferences, γ, has been fed to the decision mechanism, and that voter 1 knows this γ.

The social outcome now is a function of both γ and x. Represent this relationship as

$$P = r(x, \gamma).$$

If honesty is to be insured, whenever

$$x^* \text{ is maximizer of } r(x, \gamma)V^T, \text{ then}$$
$$x^* \equiv V.$$

Paralleling our procedure from before, with z being the utility value of the in-between alternative, represent the function r fully by

$$p_1 = j(z, \gamma),$$

and

$$p_2 = k(z, \gamma).$$

The efficiency condition that parallels [A] is for all z,

$$j_1(z, \gamma) = -zk_1(z, \gamma).^{23} \qquad [C]$$

There are no simply-stated conditions that insure the individual will write down his preferences in their true order. What is needed boils down to the tautological requirement that for all γ, the x that is the maximizer of

$$p_1 + \lambda p_2, \quad 0 \leq \lambda \leq 1$$

is of the form (1, z, 0), where z too is a value between 0 and 1.[24]

Consider a scheme that meets all requirements and is unique with respect to the individual's cardinal preferences. For such a scheme an individual with preferences (1, y, 0) will select $z=y$ for all possible values of y. From condition [C], given that z is less than 1, the increase in p_2 as z increases is greater than the corresponding decrease in p_1. The slack can only be taken up in p_3, ($p_1+p_2+p_3=1$), which therefore must be decreasing over the whole range of z. This implies that p_3 must have positive probability (except in the case when the individual is indifferent between his first and second alternatives).[25] This observation suggests another general result.

Theorem III: A voting scheme that is unique with respect to individuals' cardinal preferences must lead to probability vector outcomes all of whose elements are positive.

[22] To deal with discontinuous cases, merely replace $(f[x + \Delta x] - f[x])/\Delta x$ wherever $f'(x)$ appears, and similarly for $g'(x)$. Second derivatives can be handled in parallel fashion.

[23] This analysis could be carried out equally well for a world where other individuals' ballot markings were uncertain parameters. A subjective probability distribution $f(\gamma)$ would be placed over that parameter. Thus, efficiency condition [C] would become

$$\int j_1(z, \gamma)f(\gamma)d\gamma = -z\int k_1(z, \gamma)f(\gamma)d\gamma.$$

Other conditions would be modified equivalently; the basic results would be the same. Where there is perfect information on other individuals' preferences and ballot markings, the case considered in the text,

$$f(\gamma) = 1 \quad \text{for } \gamma = \gamma^*, \quad \text{otherwise } 0.$$

[24] One scheme that elicits true preferences for this two-individual case allows either individual to allocate 50 per cent of the outcome probability. Each individual takes half his probability as given and deals with the rest as he would in Figure 1, with the vertical scale relabeled with .5 placed where there is now a 1.

[25] If $y = z = 1$, then p_3 can be 0. As z can no longer increase, p_3 need not further decrease in value.

Shortcomings of Schemes That Are Unique with Respect to Cardinal Preferences

Theorem III has a very unfortunate implication. No democratic voting schemes that are unique with respect to individuals' cardinal preferences will guarantee Pareto-optimal outcomes. (We saw this previously when there was but one individual.) With two individuals and three alternatives, say, all three alternatives will have positive entries in the probability outcome vector. The diagram below illustrates why, in general, such a probability vector cannot represent a Pareto-optimal outcome. In the diagram, each alternative is represented by its vector of individuals' utilities.

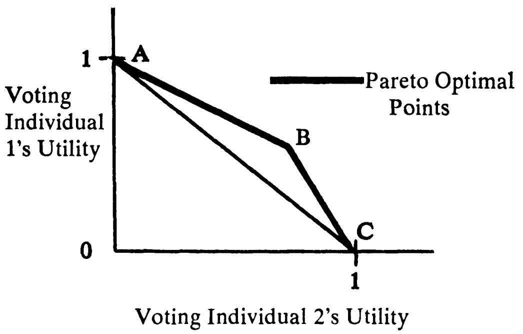


Figure 2. Pareto optimality and vectors of individuals' utilities for alternatives.

The northeast frontier represents all Pareto-optimal points. These points can be achieved through linear weighting of alternatives A and B or a linear weighting of alternatives B and C. If all three alternatives are given positive weight, a point strictly interior to the triangle ABC will be the outcome. Every such interior point is dominated by a point on the northeast frontier; therefore it cannot be Pareto-optimal. Only in the case where all three alternatives lie on the same straight line will it be possible for a Pareto-optimal scheme to give positive probability to all three alternatives. (In such a case, the concept of efficient choice loses much of its meaning. All possible probability outcome vectors are Pareto-optimal.)

The conclusion, therefore, is that in general when there are but two individuals, only two alternatives will be involved in Pareto-optimal outcomes. Converting the geometry to n dimensions, we find that with n individuals a maximum of n alternatives will need to be given positive probability weight in order to achieve all Pareto-optimal outcomes. Furthermore, assigning positive probability to more than n alternatives will lead to non-Pareto-optimal outcomes unless the utility vectors of the excess alternatives are linearly dependent (lie on the same straight line in the two-individual case) on those in some set of n.

Theorem IV: In an n-individual world, assigning positive probability to more than n alternatives whose vectors of individuals' utilities are linearly independent will result in a non-Pareto-optimal outcome.

Major Theorem

The motivation at the start of this paper was to discover a voting system that leads to Pareto-optimal outcomes. Individuals were assumed to vote in their own self-interests. It was recognized that depending on the situation and the voting system, they might or might not vote honestly. The vote-processing system would then lead to an outcome. For at least some situations we should like to require that the outcome be democratic. It was hoped that the outcome would be Pareto-optimal as well. The objective was to find systems that could guarantee these properties. Unfortunately, no such system exists.

Theorem V: No voting system that relies on individuals' self-interested balloting can guarantee a nondictatorial outcome that is Pareto-optimal.

Proof

Consider the case of two individuals with strict preferences among three alternatives whose vectors of individuals' utilities are linearly independent. By Theorem II, "if the ballot processing mechanism is to lead to Pareto-optimal outcomes, the voting individuals' ballots must be unique with respect to their cardinal preferences." By Theorem III, this voting scheme "must lead to probability vector outcomes all of whose elements are positive." But by Theorem IV, assigning positive probability to all three alternatives in this two-individual world will lead to a non-Pareto-optimal outcome. The three theorems together prove that no satisfactory voting system can be found for this case. An equivalent demonstration can be provided for any number of voters.

Summary

The recent literature of social choice has provided rather pessimistic conclusions about the possibility of passing from individuals' ordinal preferences to socially desirable outcomes. Most voting mechanisms, the most common social choice procedure where political outcomes are involved, elicit no more extensive information than individuals' ordinal preferences, and those preferences may not be honestly revealed.

This paper has confronted these problems.

First, it showed that procedures are available that will elicit individuals' honest ordinal and cardinal preferences. With honest cardinal preferences as possible inputs to a decentralized voting procedure, there was a flicker of hope for finding an efficient social choice mechanism. The flicker was extinguished when it was discovered that any social choice mechanism that elicits honest preferences automatically leads to non-Pareto-optimal outcomes. Theorem V presents the disheartening conclusion.[26] No voting system that relies on individuals' self-interested balloting can guarantee a nondictatorial Pareto-optimal outcome.

Conclusion

No decentralized mechanism can guarantee a satisfactory social choice. This result is particularly unsettling to us economists, for we are accustomed to the paradigm of the efficient market mechanism functioning in a world of perfect competition.[27]

Perhaps we should not ask despairingly, Where do we go from here?, but rather inquire, Have we been employing the appropriate mind-set all along?

We economists, to the extent that we are scientists, have fallen prey to the natural tendency to concentrate our attentions on those phenomena for which our theories have the greatest power. The land of perfect competition has been our homeground, and most times when we have ventured out, it has been to consider areas that differ in but one or two aspects from the place we know best. Our prescriptive theories for situations with externalities and public goods with rare exceptions have been developed as if the "aberrant" case we are considering is the sole deviation from the norm of the competitive model. By following this track we have achieved some useful results. But this success has only been achieved at a cost. We have constricted ourselves to view and evaluate new lands in the same terms that we have been able to apply profitably to familiar territory.

The results of this paper contribute to a muffled but growing plea: Let us not rely on market-based standards of behavior and performance when examining those important areas of the world that in no way resemble the model of perfect competition. Rather, we should describe and evaluate those areas employing terms and theories developed to meet their specific needs.

Where might this lead us? It might force us to undertake some study of the role of preference revelation. With the competitive market outcome, it is each man for himself. No one cares about anyone else's preferences, and that is the reason that no one has an incentive to disguise or distort information about his own. In all other situations, we care very much about the structure of everyone else's preferences. Self-interested people become appropriately self-conscious about revealing their preferences, and find that they have an incentive to distort them. For example, the essence of a voting problem is a common or shared decision. I shall try to influence your vote. Not only that, the way I shall vote myself will depend very much on my assessment of the voting patterns of others. I cannot be relied upon to provide the information that any vote-aggregating mechanism will wish to elicit. Decentralized decision procedures cannot be expected to function effectively. Recognizing this, we might ask, Is a fully decentralized decision procedure what we should really seek? (Indeed, to the extent that consumer sovereignty relies on individuals' expressed preferences, we might even question its universal desirability.) In political contexts, this would come down to asking, Is voting the appropriate way to make social decisions?[28]

This brings us full circle from the quote that opens this paper. That observation suggests invoking a geometric analogy for a concluding theme. Geometers may find that they can derive their most powerful results if they analyze and deal with circles, for circles are representations of perfect closed curves. But the more general category is the class of closed curves. It deserves attention too. Reference to the exceptional perfect case, the full circle in this instance, may be a type of analogy that is regularly overdrawn.

[26] New variants of Theorem IV could be developed by altering the strength of different definitions, and interchanging existential and universal modifiers. I suspect that Pareto improvements may be possible: stronger results with the same definitions, or stronger definitions leading to the same results.

[27] I am indebted to Thomas Schelling, who suggested that I emphasize this theme in my conclusion and provided me with helpful insights. This theme reiterates the central message of my volume, *Studies in Interdependence*.

[28] Many observers of the political scene would argue that our present social choice procedures are very far from aggregative processes that rely on individuals' preferences. See Edward Banfield, "'Economic' Analysis of 'Political' Phenomena; A Political Scientist's Critique" (mimeo.), 1967, and a forthcoming joint piece by Banfield and Zeckhauser on the same subject.