ERIC MASKIN
AND
AMARTYA SEN

THE ARROW IMPOSSIBILITY THEOREM

THE ARROW IMPOSSIBILITY THEOREM

KENNETH J. ARROW LECTURE SERIES

KENNETH J. ARROW LECTURE SERIES

Kenneth J. Arrow's work has shaped the course of economics for the past sixty years so deeply that, in a sense, every modern economist is his student. His ideas, style of research, and breadth of vision have been a model for generations of the boldest, most creative, and most innovative economists. His work has yielded such seminal theorems as general equilibrium, social choice, and endogenous growth, proving that simple ideas have profound effects. The Kenneth J. Arrow Lecture Series highlights economists, from Nobel laureates to groundbreaking younger scholars, whose work builds on Arrow's scholarship as well as his innovative spirit. The books in the series are an expansion of the lectures that are held in Arrow's honor at Columbia University.

Creating a Learning Society: A New Approach to Growth, Development, and Social Progress
Joseph E. Stiglitz and Bruce C. Greenwald.

THE ARROW IMPOSSIBILITY THEOREM

ERIC MASKIN | AMARTYA SEN

WITH

KENNETH J. ARROW

PARTHA DASGUPTA

PRASANTA K. PATTANAIK

JOSEPH E. STIGLITZ

COLUMBIA UNIVERSITY PRESS | NEW YORK

Columbia University Press
Publishers Since 1893
New York Chichester, West Sussex
cup.columbia.edu

Library of Congress Cataloging-in-Publication Data
Maskin, Eric.
The Arrow impossibility theorem / Eric Maskin and Amartya Sen with Kenneth J. Arrow, Partha Dasgupta, Prasanta K. Pattanaik, and Joseph E. Stiglitz.
pages cm. — (Kenneth J. Arrow lecture series)
Includes bibliographical references and index.
ISBN 978-0-231-15328-7 (cloth: alk. paper) —
ISBN 978-0-231-52686-9 (ebook)
1. Voting. 2. Social choice. 3. Probabilities. 4. Voting—Mathematical models. 5. Social choice—Mathematical models. 6. Probabilities—Mathematical models. I. Sen, Amartya, 1933- II. Arrow, Kenneth Joseph, 1921- III. Title.
JF1001.M275 2014
324.6'5—dc23
2014007031

Columbia University Press books are printed on permanent and durable acid-free paper.
This book is printed on paper with recycled content.
Printed in the United States of America

c 10 9 8 7 6 5 4 3 2 1

Cover design: Noah Arlow

CONTENTS

Acknowledgments vii

INTRODUCTION 1
PRASANTA K. PATTANAIK

PART I: THE LECTURES 23

OPENING REMARKS 25
JOSEPH E. STIGLITZ

ARROW AND THE IMPOSSIBILITY THEOREM 29
AMARTYA SEN

THE ARROW IMPOSSIBILITY THEOREM: WHERE DO WE GO FROM HERE? 43
ERIC MASKIN

COMMENTARY 57
KENNETH J. ARROW

PART II: SUPPLEMENTAL MATERIALS 65

THE INFORMATIONAL BASIS OF SOCIAL CHOICE 67
AMARTYA SEN

CONTENTS

ON THE ROBUSTNESS OF MAJORITY RULE 101

PARTHA DASGUPTA AND ERIC MASKIN

THE ORIGINS OF THE IMPOSSIBILITY THEOREM 143

KENNETH J. ARROW

Notes on Contributors 149

ACKNOWLEDGMENTS

The Kenneth J. Arrow Lecture Series has been made possible through the efforts of Columbia University's Committee on Global Thought (which I chaired when this series was inaugurated and which is now co-chaired by Saskia Sassen) and by the Program in Economic Research of the Department of Economics at Columbia University (chaired by Michael Woodford at the time of this lecture) with the support and encouragement of Columbia University Press. We are especially indebted to Robin Stephenson and Sasha de Vogel of the Committee on Global Thought and Myles Thompson and Bridget Flannery-McCoy for guiding this series to publication.

Joseph E. Stiglitz

THE ARROW IMPOSSIBILITY THEOREM

INTRODUCTION[1]

PRASANTA K. PATTANAIK

Since its publication more than six decades ago, Kenneth J. Arrow's (1950, 1951) impossibility theorem has profoundly influenced the thinking of all who are interested in issues relating to social choice and welfare, and the contributions of Eric Maskin and Amartya Sen to the vast literature, which followed Arrow's theorem, have been of fundamental importance. It is a great pleasure for me to write an introduction to this volume based on Eric Maskin's and Amartya Sen's lectures at Columbia University on Arrow's impossibility theorem, especially since, as a graduate student, I was first introduced to Arrow's impossibility theorem (and much else in welfare economics) by Amartya Sen, and I have known Eric Maskin for a long time.

WELFARE ECONOMICS AND THE THEORY OF SOCIAL CHOICE AT THE TIME OF PUBLICATION OF ARROW'S IMPOSSIBILITY THEOREM

Arrow's (1950, 1951) impossibility theorem is one of those rare intellectual contributions that virtually transform entire disciplines or subdisciplines. Welfare economics and the theory of social choice were the main areas impacted in this fashion by Arrow's theorem, though the theorem has also had far-reaching influence on political philosophy and political theory. It may be helpful to take a quick look at the state of welfare economics at the time when Arrow (1950, 1951) published his celebrated theorem. The important contributions of Bergson (1938) and Samuelson (1947), with their precise notion of a social welfare function, had already made it amply clear that economists needed to use value judgments if they wanted to engage in policy prescription or evaluation of social states, but the set of value judgments that welfare economists focused on in the 1930s and 1940s was remarkably small. The earlier utilitarian tradition was out of favor thanks to the rise of ordinal analysis in the theory of consumers' behavior and widespread skepticism about the possibility of interpersonal comparisons of utility. Much of welfare economics centered on the Pareto Principle, which says that if everybody in the society strictly prefers a social state x to another social state y, then x is strictly better than y for the society.[2] A major problem with the Pareto Principle, of course, is that, whenever any two individuals differ in their preference-based rankings over two social states, the Pareto Principle fails to compare those two social states. In an attempt to fill at least some of the numerous gaps in comparisons under the Pareto Principle, Kaldor (1939) and Hicks (1939) introduced the "compensation criterion," but it was soon clear that not only did the compensation criterion have rather limited ethical appeal but it could also sometimes yield contradictory results, such as that social state x was better for the society than social state y, and, simultaneously, social state y was also socially better than social state x. A few years before the publication of Arrow's impossibility theorem, there were several contributions[3] to the literature on social decisions through voting,[4] the papers of Bowen (1943) and Black (1948a) being particularly important. Though these contributions appeared in some of the most respected economics journals, it will perhaps be fair to say that, by and large, welfare economists did not take much interest in them until Arrow's (1950, 1951) theorem stimulated their interest in voting procedures.

ARROW'S THEOREM AND SOME INTERPRETATIONS OF THE PROBLEM OF PREFERENCE AGGREGATION

Let X be the set of all possible social states or states of affairs in the society. At any given time, the set A of actually available social states is a nonempty subset (likely a proper subset) of X. It is assumed that each individual has a preference ordering over X. Arrow's (1951) impossibility theorem deals with the problem of specifying a social ranking or ordering R of the social states in X. The social ordering R is intended to serve as the basis of social choice in the following sense: given any set A of available social states, the society chooses from A a social state that R ranks highest among the social states belonging to A.

Arrow (1951) starts with the intuition that R must be based on the profile of individual preference orderings over X (with exactly one preference ordering for each individual in the society). As Sen (2014) notes, this is clearly an intuition derived from the democratic tradition of social choice, and it underlies Arrow's (1951) formal concept of a social welfare function. Arrow's social welfare function is a function the domain of which is a class of profiles of individual orderings and the range of which is a class of social orderings. Thus, a social welfare function specifies exactly one social ordering over X for every profile of individual orderings in its domain. The concept of a social welfare function provides a formal framework for analyzing the ethical problem of how the social ordering of social states in X is to be arrived at, given the individual orderings over X. The framework would not be of much interest unless one is prepared to introduce specific restrictions or properties that one may like to impose on the social welfare function. It is, however, important to note that, intuitively, the definition of a social welfare function itself implies that cardinal information about individual preferences and interpersonal comparisons with respect to preference satisfaction cannot play any role in the determination of the social ordering. By definition, a social welfare specifies exactly one social ordering for each profile of individual preference orderings in its domain. Therefore, intuitively, if the individuals' preference orderings of social states remain the same, then the social ordering of social states must remain the same, irrespective of any changes in the cardinal information about the individuals' preferences. Arrow (1951) proposed several properties for the social welfare function. While there are alternative versions of these properties, I shall follow Sen's (2014) elegant adaptation of Arrow's properties.[5] One of the conditions that Arrow proposed was that the domain of the social welfare function should be the set of all logically possible profiles of individual orderings. Given the definition of a social welfare function, this condition amounts to the requirement that, for every logically possible profile of individual orderings, the social welfare function should specify a unique social ordering—a requirement that Sen (2014) has called the axiom of "unrestricted domain." Arrow required the social welfare function to satisfy the familiar Pareto Principle widely accepted in welfare economics. Arrow also imposed two additional properties, namely, "independence of irrelevant alternatives" and "nondictatorship," on the social welfare function (see Sen [2013] for statements of these conditions). What Arrow's impossibility theorem says is that if X contains at least three distinct social states and the society consists of a finite number of individuals, then no social welfare function can simultaneously satisfy the conditions of unrestricted domain, the Pareto Principle, independence of irrelevant alternatives, and nondictatorship. Since the qualifications that there are at least three distinct possible social states and a finite number of individuals in the society are not particularly restrictive, and since each of the four conditions seems to have some prima facie plausibility, the theorem has the flavor of a paradox. The usefulness and importance of a paradoxical result, such as Arrow's, which demonstrates that certain plausible assumptions or axioms lead to a logical contradiction, often lies in the intellectual challenge that it poses to come up with a resolution of the paradox and the resulting scrutiny of the axioms to find reasons why one should discard or modify some of them. In the case of Arrow's impossibility theorem, the challenge was tough and led to a vast and rich literature that continues to expand in various directions.

An important issue relating to Arrow's impossibility theorem involves the intuitive interpretation of the social welfare function, which, in turn, is closely linked to the interpretation of the social ordering *R* and the individual preferences that figure in the definition of a social welfare function.

One can think of at least two distinct interpretations of the social ordering. First, one can interpret *R* as reflecting the results of comparisons of social states under a procedure or rule adopted by the society to rank social states for the purpose of taking decisions; when *R* is interpreted in this fashion, the social welfare function is simply a decision procedure that the society uses to rank different social states. Alternatively, *R* can be interpreted as reflecting an individual's social welfare judgments, i.e., an individual's ethical judgments about the relative goodness or badness of social states; the individual may be an individual belonging to the society or a central planner or someone else from outside the society. The two interpretations are very different: an individual may agree to the use of a decision rule in the society that ranks two social states x and y differently from the ranking in terms of his own social welfare judgment. The distinction between the two different interpretations of the social ordering was at the center of some of the earliest assessments (see, for example, Little, 1952; and Bergson, 1954) of Arrow's theorem,[6] and it has been discussed at length by Sen (1977, 2011). Arrow (2nd edition, 1963, p. 106) accepted the validity of the distinction between the two interpretations[7] and opted for the former interpretation since he felt that the final social choice would ultimately be determined by the results of the decision procedure adopted by the society to aggregate the preference orderings of individuals.[8,9]

Like the social orderings, individual preferences also lend themselves to a variety of interpretations. Sen (2013) distinguishes three different ways of interpreting an individual's preferences: preferences as expressed through votes, preferences as reflecting the interests of the person, and preferences reflecting ethical judgments. For Arrow, an individual's preferences reflected "whatever standards he deems relevant" (see Arrow, 1951, p. 7). In many ways, Arrow's (1951) interpretation of individual preferences corresponds closely to votes. The vote cast by a voter is based on whatever the voter considers relevant: it may represent neither her interest/personal well-being exclusively nor her moral judgments exclusively.

More than anyone else, Sen has made us aware of the importance of these distinctions between the different senses in which one can talk about the social ordering and individual preferences (see Sen, 1977, 2011, 2013). The pairing of alternative interpretations of a social ordering with alternative notions of individual preferences yields various preference aggregation problems,[10] which, despite the similarity of their formal structures, are often intuitively very different. Obviously, the formal validity of Arrow's theorem does not depend on any specific interpretation of his overall formal framework. But the appeal of the framework and the appeal of specific axioms may depend significantly on the interpretation under consideration. Thus, when one interprets individual preference as the individuals' ethical assessments of social states and considers how to aggregate them to reach an ordering that will serve as the basis of the society's decisions, the fact that Arrow's definition of a social welfare function rules out cardinal aspects of individual preferences and interpersonal comparison of preference satisfaction does not seem to be a particularly restrictive feature. It is not clear that in aggregating individual moral judgments to reach social decisions one needs to compare either the intensities with which individuals consider one social state to be ethically superior to another or the individuals' levels of "moral satisfaction." Things are, however, very different when individuals' preferences are interpreted in terms of the well-being of the respective individuals and an ethical evaluator is concerned with the problem of aggregating such preferences to reach an evaluation of social states in terms of social welfare; we feel that interpersonal comparison of well-being is essential for such an exercise.

Sen (2014) has emphasized how Arrow draws on the democratic tradition in voting theory. It may be useful to distinguish here between two aspects of democracy. One is the phase of democratic deliberations, where individual judgments about alternative social states are discussed and debated; interpersonal comparisons of well-being typically play an important role in the formation of these individual judgments. As a result of deliberations and debates (or what Arrow [2014] calls "conversation and dialogue" in his commentary) in this phase, a person's initial judgments about the social states may or may not change. But a time comes when final decisions need to be taken in the society. At that stage, votes are taken to aggregate the judgments as they stand at the end of the first stage, i.e., the stage of deliberation over individual judgments. Arrow's own interpretation of his theorem was in terms of this latter phase of decision making in a democracy; it is not therefore surprising that he excluded from his formal framework the notion of cardinal preference and interpersonal comparisons of preference satisfaction.

SOME RESPONSES TO ARROW'S IMPOSSIBILITY THEOREM

One can think of two distinct types of responses to Arrow's impossibility theorem. First, one can scrutinize Arrow's overall framework and his conditions to see whether there may be a case, at least under some interpretation of the problem of preference aggregation, for modifying them, thus opening up possible routes of escape from the impossibility result. Second, one can agree that the overall framework as well as the conditions figuring in the theorem are all reasonable, at least under some interpretations of the problem of social choice, but argue that, since all preference aggregation rules are flawed insofar as none of them can possibly satisfy all those reasonable conditions, it is important to explore further the properties of these admittedly flawed rules for preference aggregation and to see whether some of them might perform better than others.

Let me briefly review some examples of the first type of response to Arrow's theorem. We have already considered how, in the context of some interpretations of the preference aggregation problem, it may be desirable to admit cardinal individual preferences and interpersonal comparisons of preference satisfaction, which are ruled out by Arrow's definition of a social welfare function. In such cases, it will also be additionally necessary to relax Arrow's condition of independence of irrelevant alternatives. Even if the definition of a social welfare function is modified to admit cardinal preferences of individuals and interpersonal comparisons of preference satisfaction, Arrow's independence of irrelevant alternatives will still rule out all such interpersonal comparisons by making the social ranking of any two social states dependent exclusively on the individuals' *rankings* of those two social states.

There are, however, other interpretations of the problem of aggregating individual preferences, where the absence of interpersonal comparison of preference satisfaction does not seem unreasonable. This seems to be the case where individual preferences are interpreted as judgments or "votes" about social states, and the social ranking is simply the result of the decision procedure adopted by the society. Let us consider Arrow's conditions in such contexts. Since the Pareto Principle and nondictatorship seem to be fairly reasonable restrictions, attempts to find a way out of Arrow's impossibility result have often concentrated on the reasonableness of other conditions. The main justification for independence of irrelevant alternatives seems to be based on convenience: if the condition is not satisfied, then, to compare any two social states, one would need information about individual preferences with respect to other ("irrelevant") social states, and this may be considered too stringent a demand for information about individual preferences.

Unrestricted domain and the definition of a social welfare function, together, imply that there must be a *social ordering* for *every logically possible* profile of individual orderings. The motivation that Arrow provided for requiring that there must be a social ordering (i.e., a binary social weak preference relation satisfying reflexivity, connectedness, and transitivity) was that, if the social weak preference relation is reflexive, connected, and transitive, then it would define a best alternative for every finite set of feasible alternatives.[11] Many rules, such as the simple majority rule, for aggregating individual preference orderings, give rise to cyclical strict social preferences, and, therefore, fail to define a best social state for some finite set of feasible social states. If the purpose of comparing social states is to make choices from different sets of feasible social states, the purpose is defeated when the social weak preference relation does not define a best social option for some sets of options. In the literature that followed Arrow (1950, 1951), it was soon noted that, while a social ordering did indeed define a best option for every finite set of feasible options, transitivity was not a necessary condition for the social weak preference relation to generate a best option for every finite set of feasible options. A number of contributions explored whether one could find a way out of Arrow's impossibility theorem by relaxing the requirement of transitivity. But it turned out that even considerably weaker restrictions than transitivity of the social weak preference relation were incompatible with other appealing properties of a preference aggregation rule. Two of the most celebrated impossibility results that used such weaker rationality conditions for the social weak preference relation were those of Gibbard (1969), who relaxed the transitivity of the social weak preference relation to transitivity of the social strict preference relation, and Sen (1970 a, b) who assumed only acyclicity of the social strict preference relation—a property that is even weaker than the transitivity of the social strict preference relation. Is there any compelling reason why one should assume that, given the individual preference orderings, social choices for every possible set of feasible social states must be based on a fixed social weak preference relation defined over the universal set of social states? The requirement that, given the individual orderings, social choices must be based on a fixed binary social weak preference relation defined over X imposes certain "consistency properties" on the society's choices. Using the notion of social choice rather than social weak preference relation as the primitive concept, Sen (1993) has shown that a counterpart of Arrow's theorem can be proved without imposing consistency properties on social choices. It can also be argued that the requirement of consistency for social choices conflicts with some of our deepest intuitions. Thus, if one agrees with John Stuart Mill (1859) that an individual should be left free to take decisions in matters involving her own "personal" life, then it is easy to demonstrate that the social states that emerge from such decision making by individuals in their respective personal spheres will sometimes violate even the weakest consistency properties for social choice discussed in the literature. Sugden (1985) gives a highly interesting example involving the choice of marriage partners, where the "social decisions" resulting from a man's freedom to propose or not to propose to a woman and a woman's freedom to accept or not to accept a man's proposal for marriage can violate even the weakest properties of choice consistency.

There is another aspect of Arrow's condition of unrestricted domain, which has been investigated at some length in the literature and to which Maskin (2014) draws our attention.[12] The condition of unrestricted domain demands that for every possible profile of individual orderings, the social weak preference relation should be an ordering. But what if some profiles of individual orderings are quite improbable? We know that the simple majority rule satisfies the Pareto Principle, independence of irrelevant alternatives, and nondictatorship. Then, it may be argued that perhaps we should not worry too much if, only for some "improbable" profiles of individual orderings, the simple majority rule yields cyclical social strict preferences, and, therefore, fails to specify a majority winner for some set of feasible alternatives. As Maskin (2014) writes, "That's the sense in which the impossibility theorem is too gloomy: if rankings are restricted in an arguably plausible way, then the five axioms are no longer collectively inconsistent." We do know of restrictions on profiles of individual orderings, which are plausible in certain contexts and which will rule out voting cycles under the majority rule. Thus, suppose the social options are political parties, all voters have a common ranking of the parties in terms of the criterion of how rightist or leftist the parties are, and the degrees of "rightism" or "leftism" of the parties are the only relevant concerns of the voters. Then we know that the individual preference orderings will satisfy Black's (1958, pp. 7–10) property of "single-peakedness," which rules out cyclicity of the social strict preference relation under the simple majority rule. But once we go beyond cases where the voters' preferences are based on the consideration of a single dimension of the options, it is difficult to think of plausible restrictions on preference profiles that will rule out such cycles under the simple majority rule. In fact, an elegant and deep result of Kramer (1973) demonstrates that, in the general case, where the options have multiple dimensions considered relevant by the voters, most restrictions on profiles of individual preference orderings, which have been discussed in the literature, are unlikely to be satisfied.

It is important to note one implication of Arrow's axioms, to which Sen (1977, 2011, 2014) has often drawn attention and which follows not from any single axiom of Arrow but from the conjunction of unrestricted domain, the Pareto Principle, and independence of irrelevant alternatives. These three axioms, together, rule out the possibility of taking into account, in the determination of the social ranking over social states, any information (e.g., information about the social states under consideration) other than information about individual preferences. Sen (2014) persuasively argues that this, by itself, constitutes a restrictive feature of a social welfare function since it conflicts with Mill's (1859) notion that a person should have the right to make his own choices with respect to matters relating to his "personal life" irrespective of how others may feel about such choice by her. The point remains valid even if we modify the definition of a social welfare function to introduce cardinal individual preferences and interpersonal comparisons of preference satisfaction. Following Sen's (1970a, 1970b) fundamental work on individual rights, a large number of writers in welfare economics and the theory of social choice have explored issues relating to individual rights, freedom, and responsibility. These contributions have extended the informational basis of social choice much beyond what is permissible under Arrow's axioms. Their intellectual origin, however, seems to lie in the intense scrutiny of the problem of social choice and social welfare evaluation that followed Arrow's impossibility theorem.

Let me now briefly comment on another type of response to Arrow's impossibility theorem. One can argue that, though every voting procedure must violate some conditions of Arrow, it is still important to know the structural properties of different voting procedures and to study whether, in some ways, some voting procedures perform better than others. While the study of individual voting procedures and their possible drawbacks goes back to Borda (1781) and Condorcet (1785), Arrow's theorem provided the intellectual stimulus for a vast number of studies of specific voting procedures and the "paradoxical" results that may arise under them.[13] These studies are of much interest. One particular line of investigation, which started with Condorcet (1985) himself, has been to calculate, on the basis of alternative assumptions, the probabilities of certain paradoxes arising under particular voting procedures (see Gehrlein and Lepelley, 2012) for an account of several results in this strand of the literature). An important study of Dasgupta and Maskin (2008), which is lucidly summarized in Maskin (2014), is concerned with the issue of how well the simple majority voting procedure performs vis-à-vis other voting procedures. The analytical framework of Dasgupta and Maskin (2008) is somewhat different from that of Arrow (1951) insofar as their analysis is presented, not in terms of social welfare functions, but in terms of what they call voting rules. A voting rule, as Dasgupta and Maskin define it, is a function, which, for every subset A, of the universal set X (assumed to be finite) of social states, and, for every profile of individual orderings over X, specifies exactly one (possibly empty) subset of A, the specified subset of A being interpreted as the set of social states that the society chooses from A, given the profile of individual orderings. It is assumed that individual preference orderings are all drawn from some nonempty subset $\mathfrak{R}$, of the set of all possible strict orderings over X. For every such $\mathfrak{R}$, five plausible properties of voting rules are defined with reference to $\mathfrak{R}$; and, for every such $\mathfrak{R}$, a voting rule is said to work well on $\mathfrak{R}$ if it satisfies all the five properties with reference to $\mathfrak{R}$. Essentially, Dasgupta and Maskin (2008) demonstrate the following:

1. If $\mathfrak{R}$ is the set of all logically possible strict preference orderings over X, then no voting rule works well on $\mathfrak{R}$ (like Arrow's impossibility theorem, this is, of course, an impossibility result, but the result is proved for Dasgupta and Maskin's voting rules rather than for social welfare functions and the axioms are intuitively somewhat different).

2. If $\mathfrak{R}$ is such that some voting rule works well on $\mathfrak{R}$, then the simple majority voting rule also works well on $\mathfrak{R}$.
3. Let V be any given voting rule of the Dasgupta-Maskin type. If certain conditions are fulfilled, then there exists a subset $\mathfrak{R}'$ of the set of all logically possible strict orderings such that the simple majority voting rule works well on $\mathfrak{R}'$ but V does not work well on $\mathfrak{R}'$.

In this specific sense, the performance of the simple majority voting rule, judged in terms of the five properties specified by Dasgupta and Maskin, dominates the performance of every other voting rule. This is so despite the fact that, when all logically possible strict individual orderings are admissible, the simple majority voting rule sometimes fails to yield a social decision for some sets of feasible social states and some profiles of individual orderings.

CONCLUDING REMARKS

Even a cursory study of the literature on the subject of social choice and social welfare evaluation reveals how deeply and extensively Arrow's impossibility theorem has influenced our thinking about every dimension of the subject. The lectures of Amartya Sen and Eric Maskin highlight, in an exceptionally lucid fashion, some major aspects of the theorem and the subsequent developments to which they themselves have contributed immensely.

NOTES

1. I am grateful to John Weymark for several helpful comments.
2. Welfare economists often use a version of the Pareto Principle, which is slightly stronger than the version given here.
3. See, for instance, Bowen (1943) and Black (1948a, 1948b, 1948c). For an integrated presentation of the material in Black's papers, as well as a history of the theory of committees and elections, see Black (1958).
4. The literature originated in the works of Borda (1781) and Condorcet (1785).
5. Maskin's (2014) versions of Arrow's restrictions are formulated for use in a framework specified in terms of social choice rather than social preference.
6. Little (1952) and Bergson (1954) felt that welfare economics should be concerned only with social welfare judgments and not with procedures or rules that the society may adopt to arrive at decisions. It is not, however, clear why welfare economics should exclude studies of the latter.
7. Arrow (2nd ed., 1963, p. 106) observes, "A welfare judgment requires that some one person is judge; a rule for arriving at social decisions may be agreed upon for reasons of convenience and necessity without its outcome being treated as evaluation by anyone in particular."
8. Arrow (2nd ed., 1963, p. 106) wrote: "'Social welfare' is related to social policy in any sensible interpretation; the welfare judgments formed by any single individual are unconnected with action and therefore sterile.'"
9. Arrow felt that his formal theorem was applicable under the second interpretation as well: "the body of welfare judgments made by a single individual are determined, in effect, by the social decision process which the individual would have society adopt if he could" (Arrow, 1963, 2nd ed., p. 106).
10. Not every such pairing may yield an intuitively interesting prob-

lem of preference aggregation. An example is the problem of an ethical evaluator of social states aggregating the individuals' ethical orderings of those social states to reach her own social welfare evaluations. While it seems highly plausible that, in reaching her own ethical evaluation of social states in terms of social welfare, the evaluator should take into account the personal well-being of all individuals in the society, it is not clear why she should take into account the moral sentiments of all those individuals.

11. Arrow (1951) also had a second justification for the requirement that the social weak preference relation should be an ordering. He argued that, if the social weak preference relation was an ordering, then the society could make its choice from a feasible set of options by using a sequence of pairwise comparisons without the final choices being dependent on the specific "path" or sequence of pairwise comparisons.
12. The central point of Maskin (2014), however, lies elsewhere, and I shall come to it a little later.
13. For a list of alternative voting rules and various voting paradoxes, see Felsenthal (2012).

REFERENCES

Arrow, K. J. (1950), "A difficulty in the concept of social welfare," *Journal of Political economy*, 58.

Arrow, K. J. (1st edition, 1951; 2nd edition, 1963), *Social Choice and Individual Values*. New York: Wiley.

Arrow, K. J. (2014), "Commentary," in *The Arrow Impossibility Theorem*. Volume 2. New York: Columbia University Press.

Bergson, A. (1938), "A reformulation of certain aspects of welfare economics," *Quarterly Journal of Economics*, 52.

Bergson, A. (1954), "On the concept of social welfare," *Quarterly Journal of Economics*, 68.

Black, D. (1948a), "On the rationale of group decision making," *Journal of Political Economy*, 1948.

Black, D. (1948b), "The decisions of a committee using a special majority," *Econometrica*, 16.

Black, D. (1948c) "The elasticity of committee decisions with an altering size of majority," *Econometrica*, 16.

Black, D. (1958), *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.

Borda, J.-C. de (1781), "Mémoire sur les élections au scrutin," *Mémoires de l'Académie Royale des Sciences*; English translation by A. de Grazia, *Isis*, 44, 1953.

Bowen, H. R. (1943), "The interpretation of voting in the allocation of economic resources," *Quarterly Journal of Economics*, 58.

Condorcet, Marquis de (1785), *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, Paris.

Dasgupta, P. and E. Maskin (2008), "On the robustness of majority rule," *Journal of the European Economic Association*, 6; also reprinted in *The Arrow Impossibility Theorem*. New York: Columbia University Press.

Felsenthal, D. S. (2012), "Review of paradoxes afflicting procedures for electing a single candidate," in D. S. Felsenthal and M. Machover (eds.), *Electoral Systems*. Heidelberg: Springer-Verlag.

Gehrlein, W. V. and D. Lepelley (2012), "The value of research based on simple assumptions about voters' preferences" in D. S. Felsenthal and M. Machover (eds.), *Electoral Systems*. Heidelberg: Springer-Verlag.

Gibbard, A. (1969), "Intransitive social indifference and Arrow's dilemma," *Review of Economic Design*, vol. 18, 3.

Hicks, J. R. (1939), "The Foundations of Welfare Economics," *Economic Journal*, 48.

Kaldor, N. (1939), "Welfare propositions in economics," *Economic Journal*, 49.

Kramer, G. (1973), "On a class of equilibrium conditions for majority rule," *Econometrica*, 41.

Lange, O. (1942), "The foundations of welfare economics," *Econometrica*, 10.

Little, I. M. D. (1952), "Social choice and individual values," *Journal of Political Economy*, 60.

Maskin, E. (2014), "The Arrow impossibility theorem: where do we go from here?" in *The Arrow Impossibility Theorem*. New York: Columbia University Press.

Mill, J. S. (1859), *On Liberty.* Reprinted in J. S. Mill, *Utilitarianism and Other Writings*, edited by M. Warnock, Penguin Books, New York, 1962.

Sen, A. (1970a), *Collective Choice and Social Welfare*. San Francisco: Holden Day.

Sen, A. (1970b), "The impossibility of a Paretian liberal," *Journal of Political Economy*, 78.

Sen, A. (1977), "Social choice theory: A re-examination," *Econometrica*, 45.

Sen, A. (1993), "Internal consistency of choice," *Econometrica*, 61.

Sen, A. (2011),"The informational basis of social choice," in K. J. Arrow, A. Sen, and K. Suzumura, *Handbook of Social Choice and Welfare*, Vol. II. Amsterdam: North-Holland.

Sen, A. (2014), "Arrow and the impossibility theorem: where do we go from here? in *The Arrow Impossibility Theorem*. New York: Columbia University Press.

Sugden, R. (1985), "Why be consistent?" *Economica*, 52.

PART I

THE LECTURES

OPENING REMARKS

JOSEPH E. STIGLITZ

The Kenneth J. Arrow lectures are given in honor of Kenneth Arrow, who was one of Columbia's most distinguished graduates. The topic of the second annual Arrow lecture, on which this book is based, was Ken's thesis *Social Choice and Individual Values*. For anyone doing a PhD, it would be no bad thing to aspire to the standards this work has set.

The fact that Ken's PhD thesis remains an icon more than a half century after its writing shows just how much it changed the way we think about the whole problem of social choice. That someone could even formulate the question his thesis poses reveals its author's novel cast of mind. I find that it is still inspiring when I read it today.

The speakers for this lecture—Amartya Sen and Eric Maskin—were particularly suitable for the occasion because of their enormous contributions to the theory of social choice, elaborating on some of the ideas and aspects of the field that Ken opened up more than fifty years ago.

Amartya Sen has been so generously frequent and welcome a visitor at Columbia that I feel we can almost claim him as one of our own. I have known Amartya for more than forty years—we met in England when I was a graduate student in the late sixties—so his presence at the lecture was a particular pleasure for me. He is now the Thomas Lamont University Professor and Professor of Economics and Philosophy at Harvard University, and until recently he was the Master of Trinity College in Cambridge.

His research has ranged over a number of fields in economics, philosophy, and decision theory, and he has made particularly important contributions to the theory of social choice, which is our broad theme today. In 1998 he received the Nobel Memorial Prize in Economics for his contribution to welfare economics, and his work on social choice was notably mentioned in the citation of the prize.

Amartya and I worked together on the Commission on the Measurement of Economic Performance in Social Progress (2009), which was set up by French president Nicholas Sarkozy to translate some of Amartya's ideas into the measurements of a country's economic performance. An earlier related work, the Human Development Index of the United Nations Development Program, has become a standard metric, especially in the context of evaluating the performance in developing countries. Amartya played a central role in creating and shaping that index.

Amartya's recent book *The Idea of Justice* is an important work that critically takes on a range of thinkers from Adam Smith to Rawls who have written on this central subject in philosophy, politics, and economics. Some of the ideas surface in his contribution to this book.

The second speaker, Eric Maskin, was the Albert Hirschman Professor at the Institute for Advanced Study at Princeton at the time of the lecture and is now the Adams University Professor at Harvard University. He is particularly well known for his work on mechanism design, including his work on how to design institutions for achieving particular social or economic goals. In recognition of his fundamental contributions, he shared the 2007 Nobel Memorial Prize in Economics.

Eric, like Amartya, has worked in many different fields. Indeed, his work has had a deep influence in almost every area of economics. It was a real pleasure to welcome him to Columbia.

ARROW AND THE IMPOSSIBILITY THEOREM[1]

AMARTYA SEN

I

It was wonderful for me to have the opportunity to pay tribute to Kenneth Arrow, who is not only one of the greatest economists of our time but also one of the finest thinkers of our era. That itself made the occasion of the second annual Arrow lecture very special for me, but on top of that, it was marvelous to have the company of Eric Maskin, with whom I used to teach a most enjoyable joint course on social choice theory at Harvard, until he deserted us for the Institute for Advanced Study at Princeton.[2] And it was very pleasing for me to have Joe Stiglitz as the *participating* chair of the meeting (having known Joe for many years, I can assure you that there was no danger of Joe being an aloof chair) and to know that Akeel Bilgrami's intellectual vision was behind the planning of this event. I was in admirable company at the lecture and want to begin this piece by expressing my appreciation of that, but most especially by thanking Ken Arrow himself, for making us all think in new lines, and personally for me, for being such a major influence on my own intellectual life.

I shall be particularly concerned in this essay with Arrow's pathbreaking "impossibility theorem," for which Arrow managed to find, in line with his sunny temperament, a rather cheerful name: "General Possibility Theorem."[3] This result, and with it the formulation of the demands of mathematical social choice theory, were real watersheds in the history of welfare economics as well as of voting theory and collective choice.

The informational foundation of modern social choice theory relates to the basic democratic conviction that social judgments and public decisions must depend, in some transparent way, on individual preferences, broadly understood. (I have investigated the implications of this perspective in social choice theory in my paper "The Informational Basis of Social Choice," which is reprinted in part 2 of this book.)[4] The emergence of this democratic instinct relates closely to the ideas and events that surrounded the European Enlightenment. Even though the pursuit of democratic social arrangements drew also on various earlier sources and inspirations, it received a definitive delineation and massive public acknowledgment only during the Enlightenment, particularly—but not exclusively—during the second half of the eighteenth century, which also saw the French Revolution and American independence.

What can be called "preferences" of persons can, of course, be variously interpreted in different democratic exercises, and the differences are well illustrated by the contrasts between (1) focusing on votes or ballots (explored in the classic works of Borda and Condorcet), (2) concentrating on the interests of individuals (explored, in different ways, in the pioneering writings of David Hume, Jeremy Bentham, and John Stuart Mill), and (3) drawing on the diverse judgments and moral sentiments of individuals about societies and collectivities (explored by Adam Smith and Immanuel Kant, among many others over the centuries). These contrasts, between alternative interpretations of preferences, can be very important for some purposes. I shall visit that territory before long. However, for the moment I shall use the generic term *preference* to cover all these different interpretations of individual concerns that could be invoked, in one way or another, to serve as the informational bases of public decisions and of social judgments.[5]

In contemporary social choice theory, pioneered by Kenneth Arrow, this democratic value is absolutely central, and the discipline has continued to be loyal to this basic informational presumption. For example, when an axiomatic structure yields the existence of a dictator as a joint implication of chosen axioms that seemed plausible enough (on this more presently), this is immediately understood as something of a major embarrassment for that set of axioms, rather than being taken to be just fine on the ground that it is a logical corollary of axioms that have been already accepted and endorsed. We cannot begin to understand the intellectual challenge involved in Arrow's impossibility theorem without coming to grips with the focus on informational inclusiveness that goes with a democratic commitment, which is deeply offended by a dictatorial procedure. This is so, even when the dictatorial result is entailed by axiomatic requirements that seem reasonable, taking each axiom on its own.

So let me begin by discussing what Arrow's impossibility theorem asserts and how it is established. The theorem has the reputation of being "formidable," which is a good description of its deeply surprising nature as well as of its vast reach, but the air of distanced respect is not particularly helpful in encouraging people to try to understand how the result emerges. It is, however, important for people interested in political science, in welfare economics, or in public policy to understand the analytical foundations of Arrow's far-reaching result, and there is no reason it should be seen as a very difficult result to comprehend and appreciate. A closer understanding is also relevant for seeing what its implications are and what alleged implications, often attributed to it, may be misleading.

The proof of the Arrow theorem I shall present follows Arrow's own line of reasoning, but through some emendations that make it agreeably short and rather easy to follow. However, it is a completely elementary proof, using nothing other than basic logic, like Arrow's own.[6] The important issue here is not just the shortness of getting to the Arrow theorem but the ease with which it can be followed by anyone without any technical reasoning or any particular knowledge of mathematics or advanced mathematical logic. So I have spelled out fully the reasoning behind each step (perhaps too elaborately for some who are very familiar with this type of logical reasoning).

II

The basic engagement of social choice with which Arrow was concerned involved evaluating and choosing from the set of available social states ($x, y, \ldots$), with each x, y, etc., describing what is happening to the individuals and the society in the respective states of affairs. Arrow was concerned with arriving at an aggregate "social ranking" R defined over the set of potentially available social states x, y, etc. With his democratic commitment, the basis of the social ranking R is taken to be the collection of individual rankings $\{R_i\}$, with any R_i standing for person i's preference ranking over the alternative social states open for social choice. It is this functional relation that Kenneth Arrow calls the "social welfare function." Given any set of individual preferences, the social welfare function determines a particular aggregate social ranking R.

That there could be problems of consistency in voting rules was demonstrated by the Marquis de Condorcet in the eighteenth century. It is useful to recollect how the problem comes about, for example, for the method of majority of decision.

Take three persons 1, 2, and 3 with the following preferences over three alternatives x, y, and z.

1	**2**	**3**
x	y	z
y	z	x
z	x	y

In majority decisions, x defeats y, which defeats z, which in turn defeats x. The R generated by majority rule violates transitivity and even weaker conditions of consistency than that (such as acyclicity). And since each alternative is defeated by another available alternative in the available set, there is no majority winner—no "choice set"—for the available set $\{x, y, z\}$.

Majority rule is of course a very special rule, though highly appealing. Arrow's theorem, among other things, generalizes the problem for any voting rule, and indeed it does much more than that (as I shall presently discuss).

Consider now the following set of axioms, which are motivated by Arrow's original axioms but are in fact somewhat simpler—and also somewhat less demanding—which, taken together, are nevertheless adequate for the impossibility theorem.

- U (*unrestricted domain*): For any logically possible set of individual preferences, there is a social ordering R.
- I (*Independence of irrelevant alternatives*): The social ranking of any pair $\{x, y\}$ will depend only on the individual rankings of x and y.
- P (*Pareto Principle*): If everyone prefers any x to any y, then x is socially preferred to y.
- D (*Nondictatorship*): There is no person i such that whenever this person prefers any x to any y, then socially x is preferred to y, no matter what others prefer.

The General Possibility Theorem: If there are at least three distinct social states and a finite number of individuals, then no social welfare function can satisfy U, I, D, and P.

One common way of putting this result is that a social welfare function that satisfies unrestricted domain, independence, and Pareto Principle has to be dictatorial. This is a repugnant conclusion—antithetical to the democratic commitment—emanating from a collection of reasonable-looking axioms.

III

In proving this theorem we can go by two intermediate results ("lemmas," if you want to sound "high-tech"). Crucial to this line of reasoning is the idea of a set of individuals, *G*, being "decisive." This could be "local" in the sense of applying over some particular pair of alternatives {*x, y*}, when *x* must be invariably socially preferred to *y*, whenever all individuals in *G* prefer *x* to *y (no matter how the individuals not in G rank x* and *y*). Much more demandingly, *G* is globally "decisive," if it is locally decisive over every pair. The first intermediate result is about what we can understand as the spread of decisiveness from one pair of alternatives to all others. Central to the strategy of demonstration used here is to postulate only a specific partial ordering for individuals not in *G*, over particular pairs (allowed by Unrestricted Domain), and the reasoning works for every complete ordering of individuals compatible with that partial ordering.

SPREAD OF DECISIVENESS

If G is decisive over any pair {x, y}, then G is [globally] decisive.

Proof: Take any pair {*a, b*} different from {*x, y*}, and assume that everyone, *without exception,* prefers *a* to *x*, and *y* to *b.*

For those who do not belong to G, we do not impose any other condition on the rest of their preferences (in particular they can rank a and b in any possible way). But we assume that all members of group G also prefer x over y; that is, they subscribe to the descending order: a, x, y, b. By the Pareto Principle, a is socially preferred to x, and y is socially preferred to b. By the decisiveness of G over $\{x, y\}$, x is socially preferred to y. Putting them together (that is, a preferred to x, that to y, and that to b), we have, by transitivity of strict preference, the result that a is socially preferred to b. By the Independence of Irrelevant Alternatives (condition I), this must relate only to individual preferences over $\{a, b\}$. But only the preferences of individuals in G have been specified (they rank a above b); all others can rank a and b in any way they would like. So G is also decisive over the pair $\{a, b\}$, and not just over $\{x, y\}$. And this applies to all pairs $\{a, b\}$ distinct from $\{x, y\}$. So G is indeed [globally] decisive.[7]

CONTRACTION OF DECISIVE SETS

If a set G of individuals is decisive (and if it has more than one individual), then some reduced part (a "proper subset") of G is decisive as well.

Proof: Partition G into two subsets G_1 and G_2. Let everyone in G_1 prefer x to y, and x to z, with the ranking of y and z unspecified, and let everyone in G_2 prefer x to y, and z to y. Others not in G can have any set of preferences. By the decisiveness of G, we have x socially preferred to y. If now z is taken to be socially at least as good as x for some configuration of individual preferences over $\{z, x\}$, then we must have z socially preferred to y (since x is socially preferred to y) for that configuration of preferences over $\{z, x\}$. Since no one's preference over $\{z, y\}$ other than those in G_2 has been specified,[8] and those in G_2 prefer z to y, G_2 is decisive over $\{z, y\}$, and thus, by the Spread of Decisiveness, G_2 must be [globally] decisive. Since that shows that some reduced part of G is indeed decisive, then we have got what we want to show, for that particular case. To avoid this possibility, we must assume that our initial presupposition that z is at least as good as x must be eschewed. But then x must be preferred to z. However, since no one's preference over $\{x, z\}$ has been specified for this, other than those in G_1 who prefer x to z, clearly G_1 is decisive over $\{x, z\}$. Thus by the Spread of Decisiveness, G_1 is [globally] decisive. So either G_1 or G_2 must be [globally] decisive, which establishes the Contraction of Decisive Sets.

Now, Arrow's impossibility result.

PROOF OF THE GENERAL POSSIBILITY THEOREM

By the Pareto Principle the set of all individuals is decisive. By the Contraction of Decisive Sets, some proper subset of all individuals must also be decisive. Take that smaller decisive set, but some proper subset of that smaller set must also be decisive. And so on. Since the set of individuals is finite, we shall arrive, sooner or later, at one individual who is decisive. But that violates the nondictatorship condition—hence the impossibility.

IV

I end with a few observations on the nature of Arrow's result. First, the combination of unrestricted domain, independence, and the Pareto Principle—each of which is individually somewhat innocuous—seems to produce both the spread of decisiveness and the contractibility of decisive sets. One lesson to draw is that it is hard to judge the plausibility of axioms unless we also consider with what other axioms they are to be harnessed together.

Second, even though the basic axioms for the Arrow theorem concentrate on individual preferences over the set of alternative social states, it is not directly presumed by any of the axioms that we are not allowed to take note of the nature of the alternatives and their relations to the persons involved in arriving at the social ranking R. That ruling out of information about the nature of the alternatives results from the combination of the three axioms—U, I, and P—that together entail Spread of Decisiveness, ruling out the relevance of the nature of the alternatives involved in the choice, and of the comparative predicaments of different individuals in different states of affairs. This is what causes the permitted social welfare functions to be confined to the class of voting rules. So, in this sense, it is wrong to think of the Arrow result as merely extending the Condorcet paradox to all voting rules. It first establishes that the permitted social welfare functions must be voting rules (that is the big intermediate result), and *then* generalizes the Condorcet paradox.

Third, the result can be easily extended to social choice reasoning that avoids talking about any social relation *R* and concentrates instead on what can be chosen from particular sets of alternatives. This is called, in the literature, "social choice functions." The ultimate choice-functional extension is to impose no internal consistency conditions on social choice functions at all, except those that come from the relation between individual preferences and the social choice function. To give an example, the choice-functional version of the Pareto Principle would be: If everyone prefers *x* to *y*, then *y* must not be chosen if *x* is available to be chosen. The Arrow theorem can indeed be extended to choice functions as well, without any internal consistency condition, as I have shown elsewhere (but, I fear, that proof is too complicated to be aired here!).[9]

Fourth, the individual preferences are just the orderings of individuals considered separately, without any interpersonal comparisons. Once interpersonal comparisons are allowed, various positive possibilities open up. This is important for welfare economics and for judgments of equity and aggregate welfare with which welfare economists are rightly concerned. Consider a nasty proposal to take some of the income of the poorest person and divide it over several others. In a society of selfish persons, this unprepossessing proposal will be a majority improvement. So the problem here is not the lack of consistency of majority rule or any other voting rule. It is that we are in wrong territory by concentrating only individual preference orderings, and then—with the help of combining *U*, *I*, and

P—getting to the Spread of Decisiveness. It is the wrong informational base for many welfare economic concerns, and it is perhaps all to the good that the majority rule is also—in addition to being obtuse—inconsistent.

Fifth, the way to tackle the Arrow theorem in the context of welfare economics certainly includes making use of interpersonal comparisons in our judgments. Indeed, all public policy tends to bring in interpersonal comparisons in one way or another. But that route is not easily available when we are dealing with a voting process, such as elections of candidates for, say, political positions. For the political exercise and voting theory, we have to think in different lines, as indeed Eric Maskin does, with his usual skill and elegance, in his own lecture here.[10]

Finally, even for political processes, one problem that remains does not involve interpersonal comparisons but for which the Arrovian axioms are inadequate. This is the problem of liberty and rights. If we follow John Stuart Mill in standing up for the rights of minorities and of individuals in their personal domain, then we must not be too impressed by how many people oppose the minorities being able to choose their own life styles or how many try to eliminate the exercise of personal liberties by individuals. The importance of liberty may demand going against the objections of numerous "busybodies." Not to be able to accommodate liberty is a political limitation of majority rule. It is not accommodated in Arrow's axioms either.

All these problems, and many more, open up as we follow—and follow up on—Arrow's pioneering result, including his demonstration that serious problems can arise with axiomatic combinations even when each individual axiom, seen on its own, seems just fine. That was indeed a game-changing contribution, which affected a large variety of disciplines. It is difficult to think that Arrow was just a graduate student when he transformed the intellectual world of social thought.

NOTES

1. I am grateful for the comments on my presentation at the meeting at Columbia, including those from Eric Maskin and Joseph Stiglitz.
2. Happily, Maskin rejoined the faculty at Harvard in 2012.
3. Kenneth J. Arrow, *Social Choice and Individual Values* (New York: Wiley, 1951; republished in an extended form, 1963).
4. Amartya Sen, "The Informational Basis of Social Choice," in *The Handbook of Social Choice and Welfare,* vol 2, eds. Kenneth Arrow, Amartya Sen, and Kotaro Suzumura (Amsterdam: Elsevier, 2010).
5. These distinctions and their far-reaching implications were investigated in my essay "Social Choice Theory: A Re-examination," *Econometrica* 45 (1977); reprinted in *Choice, Welfare and Measurement* (Oxford: Blackwell, and Cambridge, Mass.: MIT Press, 1982; republished, Cambridge, Mass.: Harvard University Press, 1997), and further in my Presidential Address to the American Economic Association: "Rationality and Social Choice," *American Economic Review* 85 (1995); reprinted in *Rationality and Freedom* (Cambridge, Mass.: Harvard University Press, 2002).
6. Providing short proofs of Arrow's theorem is something of a recurrent exercise in social choice theory, and one must not make a cult of it, since all the proofs draw in one way or another on Arrow's trail-blazing insight. One has to be careful, however, not to be too convinced by the alleged "shortness" of some proofs when they draw on other mathematical results (as some of the

proofs do) that are invoked and used but not established in the course of the proof. Arrow's proof was entirely elementary and did not presume anything other than the basic rules of logic, which is the route followed here. I presented an earlier version of this proof in footnotes 9 and 10, in "Rationality and Social Choice" (1995); reprinted in *Rationality and Freedom* (2002), p. 267.

7. We are cutting a corner here by assuming that a, b, x, and y are all distinct social states. The reasoning is exactly similar when two of them are the same alternative.
8. Note that this lack of specification of individual preferences over $\{z, y\}$ is consistent with what has been assumed about individual preferences over $\{z, x\}$.
9. See my Presidential Address to the Econometric Society, "Internal Consistency of Choice," *Econometrica* 61 (1993); reprinted in *Rationality and Freedom* (2002).
10. Eric Maskin, "The Arrow Impossibility Theorem: Where Do We Go from Here?" in this volume.

THE ARROW IMPOSSIBILITY THEOREM: WHERE DO WE GO FROM HERE?[1]

ERIC MASKIN

Giving a lecture in honor of Kenneth Arrow would be a high point for any economist, but, in my case, there were two additional reasons why the occasion of the second annual Arrow lecture was such a special pleasure.

First, Ken Arrow was my teacher and PhD advisor, and most likely I would not have become an economist at all had it not been for him. I was a math major in college and intended to continue in that direction until I happened to take a course of Ken's—not on social choice theory but on "information economics." The course was a hodgepodge—essentially, anything that Ken felt like talking about. And it often seemed as though he decided on what to talk about on his way to the classroom (if then); the lectures had an improvised quality. But they were mesmerizing, and, mainly because of that course, I switched to economics.

Second, lecturing with Amartya Sen brought back many happy memories for me, because he and I taught this book's subject—social choice theory—together several times in a graduate course at Harvard. It's great to renew our pedagogical partnership.

Like Amartya, my essay is on the Arrow impossibility theorem, but I will concentrate on its implications for voting and elections; I will leave aside its broader implications for social welfare.

Now, by its very name, the impossibility theorem engenders a certain degree of pessimism: If something is "impossible," it's pretty hard to do. As applied to voting, the theorem appears to say that there is no good election method. Well, I will make the case that this is too strong a conclusion to draw; it's overly negative. But whether or not I persuade you of this, I want to argue that the theorem inspires a natural follow-up question, which oddly was not addressed until quite recently. And I will discuss that question and its answer at the end of this talk.

Let me begin by reviewing the impossibility theorem from the standpoint of elections. If there is a political office to fill, then a *voting rule* is a method of choosing the winner from a set of candidates (this set is called the *ballot*) on the basis of voters' rankings of those candidates.

Many different voting rules have been considered in theory and practice. Probably the most widely used method here in the United States is *plurality rule*, according to which the winner is the candidate who is more voters' *favorite candidate* (i.e., the candidate more voters rank first) than any other.[2] Thus, if there are three candidates *X*, *Y*, and *Z*, and 40 percent of the electorate like *X* best (i.e., 40 percent rank him first), 35 percent like *Y* best, and 25 percent like *Z* best (see table 1), then *X* wins because 40 percent is bigger than 35 percent and 25 percent—even though it is short of an overall majority. Plurality rule is the method used to

Table 1 *X* is the plurality winner

40%	35%	25%
X	*Y*	*Z*

elect senators and representatives in the United States and members of parliament in Britain (where it's called "first-past-the-post.")

Another well-known method is *simple majority rule*,[3] which the eighteenth-century French mathematician and philosopher Condorcet was the first to analyze in detail.[4] The winner under majority rule is the candidate who is preferred by a majority to each other candidate. For instance, suppose there are again three candidates, *X*, *Y*, and *Z*. Forty percent of voters rank *X* first, then *Y*, and then *Z*; 35 percent rank *Y* first, then *Z*, and then *X*; and 25 percent rank *Z* first, then *Y*, and then *X* (see table 2). Based on these rankings, the majority winner is candidate *Y*, because a majority of voters (35 percent + 25 percent = 60 percent prefer *Y* to *X*, and a majority (40 percent + 35 percent = 75 percent) prefer *Y* to *Z*.

Table 2 *Y* is the plurality winner

40%	35%	25%
X	*Y*	*Z*
Y	*Z*	*Y*
Z	*X*	*X*

Notice that plurality rule and majority rule lead to different outcomes: For the voter rankings of table 2, plurality rule elects candidate *X*, whereas majority rule chooses *Y*. This difference prompts an obvious question: Which outcome is "right"? Or, put another way, which voting rule is better to use? Indeed, there is no reason to stop with plurality or majority rule; we can ask which among all *possible* voting rules is best.

Arrow provided a framework for answering these questions. He proposed that we should first try to articulate what we *want* out of a voting rule, that is, what properties—or *axioms*—we want it to satisfy. The best voting rule will then be the one(s) that fulfill all those axioms.

Here are the axioms that Arrow considered. As we will see, each is highly desirable on its own, but collectively they lead to impossibility. Because I am particularly concerned with elections, I will suggest reformulations of the axioms that are particularly suited to such contests.

The first is the requirement that an election be *decisive*, i.e., that, whatever voters' rankings turn out to be, there should always be a winner and there shouldn't be more than one winner.[5] The second is what an economist would call the *Pareto Principle* and what a political theorist might call the *consensus* principle: the idea that if all voters rank candidate *X* above candidate *Y* and *X* is on the ballot (so that *X* is actually available), then we ought not to elect *Y*. The third axiom is the requirement of *nondictatorship*—no voter should have the power to always get his way. That is, it should not be the case that if he likes candidate *X* best, then *X* is necessarily elected, regardless of how others feel about *X*. Otherwise, that voter would be a dictator.

The final Arrow axiom is called *independence of irrelevant alternatives*, which in our election context could be renamed "independence of irrelevant candidates." Suppose that, given the voting rule and voters' rankings, candidate *X* ends up the winner of an election. Now look at another situation that is exactly the same except that some other candidate *Y*—who didn't win—is no longer on the ballot. Well, candidate *Y* is, in a sense, "irrelevant"; he didn't win the election in the first place, and so leaving him off the ballot shouldn't make any difference. And so, the independence axiom requires that *X* should still win in this other situation.[6]

I think that, put like this, independence seems pretty reasonable, but its most vivid justification probably comes from actual political history. So, for example, let's recall the U.S. presidential election of 2000. You may remember that in that election everything came down to Florida. If George W. Bush carried the state, he would become president and the same for Al Gore. Now, Florida—like most other states[7]—uses plurality rule to determine the winner. In the event, Bush got somewhat fewer than six hundred more votes than Gore. Although this was an extraordinarily slim margin in view of the nearly 6 million votes cast, it gave Bush a plurality (and thus the presidency). And, leaving aside the accuracy of the totals themselves (hanging chads and the like), we might reasonably ask whether there was anything wrong with this outcome. But the answer to this question is made more complicated by the presence of a third candidate in Florida, Ralph Nader. Nearly one hundred thousand Floridians voted for Nader, and it is likely that, had he not been on the ballot, a large majority of these voters would have voted for Gore (of course, some of them might not have voted at all). That means that Gore would probably not only have won, but won quite handily, if Nader had not run.

In political argot, Nader was a *spoiler*. Although he got less than 2 percent of the vote in Florida—he was clearly "irrelevant" in the sense of having no chance to win himself—he ended up determining the outcome of the election. That seems highly undemocratic.[8]

The independence axiom serves to rule out spoilers. Thus, because plurality rule was quite spectacularly vulnerable to spoilers, we can immediately conclude that it violates independence. Majority rule, by contrast, is easily seen to satisfy independence: if candidate X beats each other candidate by a majority, it continues to do so if one of those other candidates is dropped from the ballot.

Unfortunately, majority rule violates our first axiom, decisiveness—it doesn't always produce a clear-cut winner (this is a problem that Condorcet himself discussed). To see what can go wrong, consider an election with three candidates X, Y, and Z, and an electorate in which 35 percent of the population rank X first, Y second, and Z third; 33 percent rank Y first, Z second, and X third; and 32 percent rank Z first, X second, and Y third (see table 3).

Table 3 *Y* is the plurality winner

35%	33%	32%
X	Y	Z
Y	Z	X
Z	X	Y

Observe that *Y* beats *Z* by a majority (68 percent to 32 percent), and *X* beats *Y* by a majority (67 percent to 33 percent). But *Z* beats *X* by a majority (65 percent to 35 percent)—and so there is no candidate who beats each of the other two. This phenomenon is called the Condorcet paradox.

Interestingly, Kenneth Arrow wasn't aware of Condorcet's paradox when he started work on social choice theory. He rediscovered it while studying how firms might make choices. In economics textbooks, firms choose production plans to maximize their profit. But in reality, of course, a firm is not typically a unitary decision maker; it's owned by a group of shareholders. And even if every shareholder wants to maximize profit, different shareholders might have different beliefs about which production plans will accomplish that. So, there has to be a choice method—a voting rule—for selecting the actual production plan.

Ken's first thought was to look at majority rule as the method, but soon discovered—or, rather rediscovered—the Condorcet paradox. Now, he knew that majority rule had been around for a long time and so assumed that his discovery couldn't possibly be novel. Indeed, when he wrote it up, he referred to it as the "well-known" paradox of voting. It was only after publication that readers directed him to Condorcet.

Although majority rule violates decisiveness and plurality rule violates independence, Ken felt that surely there must be other voting rules that satisfy all four axioms: decisiveness, consensus, nondictatorship, and independence. But after trying out rule after rule, he eventually came to suspect that these axioms are collectively contradictory. And that's how the impossibility theorem was born; Ken showed that there is no voting rule that satisfies all four axioms.[9]

Now, the nondictatorship axiom is very undemanding. For instance, if instead of one voter, *two* voters out of the entire electorate have all the power in determining the winner, we probably still won't be terribly happy with the election method, even though *nondictatorship* will then formally be satisfied. Democratic societies usually insist on the stronger condition *equal treatment of voters*, the requirement that all voters have the same weight. Equal treatment of voters is called *anonymity* in voting theory, reflecting the idea that voters' *names* shouldn't matter; only their *votes* should. Indeed, just as we require that voters be treated equally, we ordinarily do the same for candidates too: we demand *equal treatment of candidates* (called neutrality in the voting theory literature). But because Arrow showed that impossibility results from requiring decisiveness, consensus, independence, and nondictatorship, we get impossibility a fortiori from imposing the more demanding set of axioms: decisiveness, consensus, independence, equal treatment of voters, and equal treatment of candidates.

The impossibility theorem has been the source of much gloom because, individually, each of these five axioms seems so compelling. But, as I suggested in my opening remarks, there is a sense in which the theorem overstates the negative case. Specifically, it insists that a voting rule satisfy the five axioms *whatever* voters' rankings turn out to be (i.e., for an unrestricted domain of rankings). Yet, in practice, some rankings may not be terribly likely to occur. And if that's the case, then perhaps we shouldn't worry too much if the voting rule fails to satisfy all the axioms for those improbable rankings.

For an example, let's go back to the U.S. presidential election of 2000. The three candidates of note were Bush, Gore, and Nader. Now, many people ranked Bush first. But the available evidence suggests that few of these voters ranked Nader second. Similarly, a small but significant fraction of voters placed Nader first. But Nader aficionados were very unlikely to rank Bush second.

Indeed, there is a good reason why the rankings Bush/Nader/Gore (Bush ranked above Nader and Nader ranked above Gore) or Nader/Bush/Gore appeared to be so rare. In ideological terms, Nader was the left-wing candidate, Bush was the right-wing candidate, and Gore was somewhere in between. So, if you liked Bush's proposed policies, you were likely to revile Nader's, and vice versa.

Yet, if we can rule out the two rankings above (or, at least, assign them low enough probability), then it turns out that the Condorcet paradox cannot occur, and majority rule is decisive after all—it always results in a clear-cut winner. That is, majority rule satisfies all five axioms—decisiveness, consensus, no spoilers, and the two equal treatment properties—when the six logically possible rankings of Gore, Bush, and Nader are *restricted* to rule out the rankings Bush/Nader/Gore and Nader/Bush/Gore.

That's the sense in which the impossibility theorem is too gloomy—if rankings are restricted in an arguably plausible way, then the five axioms are no longer collectively inconsistent. But regardless of whether you accept the plausibility of this particular restriction, the impossibility theorem prompts a natural follow-up question: Given that no voting rule satisfies the five axioms all the time, which rule satisfies them *most often*? In other words, if we can't achieve the ideal, which voting rule gets us closest to that ideal and maximizes the chance that the properties we want are satisfied?

Perhaps, surprisingly, this question doesn't seem to have been formally posed in the literature until many years after the publication of *Social Choice and Individual Values*.[10] (Note that a paper I did with Partha Dasgutpa that addresses this question is reprinted in part 2 of this book.) In an effort to provide an answer, let me define that a voting rule works well if, for a particular restricted class of rankings, it satisfies the five axioms whenever voters' rankings adhere to the restriction. So, for example, majority rule works well in the U.S. presidential election example if rankings are restricted to exclude the two rankings Bush/Nader/Gore and Nader/Bush/Gore.[11] The question then becomes: What is the voting rule that works well for as many different restricted classes of rankings as possible?

It turns out that there is a sharp answer to this problem, provided by a "domination theorem."[12] The theorem can be expressed as follows. Take any voting rule that differs from majority rule, and suppose that it works well for a particular class of rankings. Then, majority rule must also work well for that class. Furthermore, there must be some other class of rankings for which majority rule works well and the voting method we started with does not. In other words, majority rule dominates every other voter rule: whenever another voting rule works well, majority rule must work well too, and there will be cases where majority rule works well and the other voting rule does not.[13]

As mentioned in footnote 2, I have been assuming that voters report their rankings *sincerely*—that there is no difference between their expressed rankings and their true rankings. But for many voting systems, strategic voting (ranking A above B even if you prefer B to A) may at times be advantageous. Nevertheless, the domination theorem I stated goes through unchanged if we add an additional axiom, *strategy-proofness* (the requirement that voters should find it in their interest to vote sincerely), to the list. Indeed, if majority rule satisfies decisiveness for a class of rankings, it also satisfies strategy-proofness for that same class.

I noted before that Kenneth Arrow himself began with majority rule when he set off on his examination of social choice theory. He was soon led to consider many other possible voting rules too. But it turns out that, using the criteria he laid out, there is a sense in which we can't do better than majority rule after all.

NOTES

1. I thank Amartya Sen and Joseph Stiglitz for helpful comments on the oral presentation of this lecture. The NSF provided research support.
2. Here I make no distinction between a voter's *expressed* ranking of the candidates and his *actual* ranking. In other words, I suppose voters vote "sincerely": If a voter says that candidate X is his favorite, then X really is his favorite. Still, the possibility that

voters might vote differently from their actual rankings—i.e., vote strategically—is an interesting and realistic possibility that I will return to at the end of the lecture.

3. There are many variants of majority rule, but here I will consider only Condorcet's version. I shall, therefore, omit the word *simple* from now on.
4. See M. J. Condorcet, *Essai sur l'application de l'analyse à la pluralité des voix* (Imprimerie Royale, 1785).
5. In Arrow's own framework, decisiveness is formulated as *transitivity*, the requirement that if, given voters' rankings, candidate *X* is elected over candidate *Y* and *Y* is elected over *Z*, then *X* should be elected over *Z*. (Note that if transitivity is not satisfied—so *Z* is elected over *X*—and *X*, *Y*, and *Z* are all on the ballot, then *none* of them is elected, a failure of decisiveness.)
6. There are two closely related conditions that go under the name "independence of irrelevant alternatives": Arrow's axiom and the condition formulated by J. Nash ("The Bargaining Problem," *Econometrica* 18(2), 1950: 155–162). Here I am using the Nash formulation, because it is somewhat more convenient for my purposes.
7. All other states, in fact, except, currently, for Maine and Nebraska.
8. I have singled out the election of 2000. But the same problem has recurred many times in U.S. presidential election history. For example, Ross Perot may well have spoiled the 1992 election for George H. W. Bush, enabling Bill Clinton to win.
9. In fact, Arrow also imposed a fifth axiom, *unrestricted domain of preferences*: the requirement that a voting role be defined regardless of what voters' preferences turn out to be. This axiom is implicit in the way I have formulated the other axioms.
10. See E. Maskin, "Majority Rule, Social Welfare Functions, and Games Forms," in *Choice, Welfare, and Development* (Essays in Honor of Amartya Sen), eds. K. Basu, P. Pattanaik, and K. Suzumura (Oxford: Oxford University Press, 1995), 100–109; and P. Dasgupta and E. Maskin, "On the Robustness of

Majority Rule," *Journal of the European Economic Association*, 2008: 6: 949–973.

11. But that is not the only case in which majority rule works well—excluding Gore/Nader/Bush and Bush/Nader/Gore, for example, would also do.
12. See the Dasgupta/Maskin article.
13. Of course, if we are to use majority rule in practice, we must have a tie-breaking rule in place in case the rankings of the Condorcet paradox arise—even if they are unlikely to do so. One possibility is to use plurality rule to break the tie. So, to return to the example of table 3—where candidates *X*, *Y*, and *Z* are essentially tied according to majority rule—we would then choose *X*, who enjoys the highest plurality, as the winner.

COMMENTARY[1]

KENNETH J. ARROW

I am very grateful to the two lecturers; it was indeed an honor to have Eric Maskin and Amartya Sen speak about my now quite old impossibility theorem. I cannot imagine two better discussants. Let me first turn to Eric's presentation. He presented an extraordinarily interesting theorem about the situations in which the impossibility theorem fails. In other words, he imposes restrictions on the sets of individual preferences and finds that under his conditions majority voting will work. Of course he added the condition of anonymity to the ones I impose in order to achieve that result. It is an extraordinary simple equation, and he puts great emphasis on the role of majority voting in the sense of Condorcet. That is to say, we consider all the pairwise comparisons and pick the candidate who beats everyone individually.

However, there is of course one condition that does seem to be essential to any kind of social choice rule—namely, that it give a result. That is to say, in the usual terminology, that it be decisive. I do not yet quite understand how Eric's results can help us in the case where his conditions fail.

Something has to happen if majority voting is intransitive or, in other words, where the restrictive set of preferences is insufficient to overcome the impossibility theorem. This is a pretty key issue. So I leave that as an open question for Eric to analyze further.

I will now turn to Amartya's insights and beautiful exposition of my theorem. I fully agree with his comments, but I have two remarks. First, he notes that I called the result a possibility theorem and attributes this to my sunny disposition. The facts are a little different. I have always regarded myself rather as a gloomy realist—but perhaps I am wrong. Instead, someone else, Tjalling Koopmans,[2] insisted on using the word *possibility*. He was upset by the term *impossibility*. Now, I cannot say that Tjalling had an extraordinarily sunny disposition, either. He was not necessarily a lively or cheerful person, nor was he really an optimist. But he did dislike the feeling that things could not happen or change. And given that the dissertation was originally posed as a Cowles Commission monograph, I felt that to please Tjalling, I would call it a "possibility theorem." It was not, however, my idea at all.[3]

Another issue concerns Amartya's discussion of the informational basis of preference in relation to interpersonal comparisons. He notes that in my original formulation the orderings of individuals were considered separately without any interpersonal comparisons. And Amartya is correct that I did not address the question of conversation and dialogue in the formation of preferences, the meaning of changing your opinion when talking to someone else. That said, the contention that the impossibility theorem ignores interpersonal comparisons slightly misinterprets my intent. While I did not allow for the interpersonal comparison of utilities, this did not mean that interpersonal comparisons were completely excluded. Instead, the individual preferences that form the basis of the social welfare function involve such comparisons. Individual preferences are about society. Of course the individual might give more weight to themselves than others, but the pairwise comparisons are rankings of different social orderings. Thus in constructing the impossibility theorem, a particular type of interpersonal comparison was present.

The organizers of the lecture and this volume have asked me to address a few other items. One concerns what aspects of social choice theory I would be interested in pursuing today. From a technical point of view, I would like to see more research on the condition that I find, in a way, to be the most problematic: the independence of irrelevant alternatives. In other words, if this condition is relaxed is there a decision mechanism that will satisfy the other restrictions? I might, for instance, put in a few candidates on the ballot who were not actually available. They could serve, in a way, to give some measurement to the others. Now I do not exactly know how to do this in a way that is going to be consistent. But by dropping the independence of irrelevant alternatives condition, you could then employ a Borda count, the method proposed by Balinski and Laraki,[4] or perhaps some other alternative. Regardless, what emerges is a consistent ranking, which satisfies the other conditions, not based on cardinal utilities but instead on the rankings. Of course if a candidate drops out, different results emerge. Whether this is a good decision mechanism or not, I am not prepared to say. Regardless, it is an interesting area for further exploration.

Related to the relaxation of the condition of irrelevant alternatives are efforts to develop an interpersonal scale. If you look at the questionnaires that economists use to assess, for instance, the happiness of people, they will ask a respondent, "How happy are you?" The person then ranks their happiness on a scale. Now a rock bottom, hard-boiled economist might say to the individual, "You had a happiness of one, and now it's three? What do you mean?" Such an approach is somewhat equivalent to the work of the psychologist Stanley Smith Stevens whose research had people make comparisons that were equivalent to asking, "Is this light brighter than this sound is loud?"[5] Instead of saying, "I don't know what you are talking about," people respond to these questions and find them meaningful. And if people find them meaningful, then I have to say that in my point of view they must be so. The responses of individuals permit, moreover, a systematic representation. Thus the work of economists, such as Daniel Kahneman and Alan Krueger and their "Day Reconstruction Method,"[6] is more or less equivalent to the development of an interpersonal scale and overcoming the independence of irrelevant alternatives. While I do not have a mathematical theorem capturing the findings of these studies, the formulation of some other kind of condition based on this research that replaces the independence of irrelevant alternatives would be an interesting avenue to explore.

Another and somewhat related area, and one that has important relevance to climate change issues, is the comparison of utility streams over time and what special properties we want associated with such comparisons. Some people try to argue for complete symmetry or that people a few hundred years from now are equal to individuals today—in other words, the principle of universalizability. Such a stance leads, however, to a paradox of sorts if you go out to infinity. If we impose the condition, for example, that if one stream is better in every period than another, or at least as good in every period and better in one, it should be better. But if we are all alike, and I change, would the second person or the individual next to them, and so on down to the person in a hundred years also change so that you could switch them around? It should make no difference, yet it does. Under the condition of universalizablity, if future individuals are going to be better off than we are, then our willingness to sacrifice on their behalf is certainly reduced.[7] Such paradoxes stress the central importance of the axioms that we utilize when thinking about utility streams over time and issues of intergenerational equity in relation to climate change.

Related to the previous matter is the issue of constructability. When you are dealing with infinite dimensional elements, can you really compute the results? Some things are simply quite extremely difficult to compute. They're not constructible in the sense that there is no finite process that will enable an individual to carry out the calculation. This applies to a lot of problems, not just those that are social in nature, such as climate change, but also to individual as well as social choice problems. To put it more simply, you could say, "You choose the best of that heap." But then how one exactly does that can be quite complicated if not impossible in a finite length of time.

While the points just made are focused on future research paths related to the impossibility theorem, the organizers asked me to speak about the various inspirations for my dissertation. Many of these sources have been discussed at length elsewhere, but I do have a related anecdote to end with that speaks to the, shall we say, unexpected sources of ideas. I cannot say that my interest in elections was extraordinary and that it steered me toward the impossibility theorem. Yet I do distinctly recall an incident when I was eleven years old. It revolved around the 1932 Democratic Convention. In those days the conventions were a big deal. There were a lot of candidates, and at the time only a handful of states had primaries. Instead the local political machines would designate their delegates, and the presidential candidate was selected at the convention (that year the New York delegation was predominantly for Alfred Smith and not for Franklin Delano Roosevelt). The convention really decided things, and so, like most people, I listened. I remember turning on the radio and getting my sister, who was seven years old. I made a big chart with all the candidates on the left and all the states along the top or something to that effect. A lot of the states had their own local candidate who would be run for several ballots, just to get the name nationally known. During the voting every state reported and their selections were called out, such as, "Alabama casts 24 votes for Alfred Smith" and so and so. I then dictated to my sister what to write down on the chart during all the rounds of voting. To this day she teases me that given her assistance in this endeavor that she was the one who started me on my social choice career!

NOTES

1. I would like to thank Rachel Harvey, PhD, for her editorial assistance in compiling these comments.
2. Tjalling Charles Koopmans (1910–1985): MA in physics and mathematics, University of Utrecht, 1933; PhD in mathematical statistics, University of Leiden, 1936; lecturer, Netherlands School of Economics, 1936–1938; economist, League of Nations, Geneva, 1938–1940; research associate, Princeton, 1940–1941; Penn Mutual Life Insurance Co., 1941–1942; statistician, Combined Shipping Adjustment Board, 1942–1944; Cowles Commission, 1944–1945; professor, University of Chicago, 1946–1955; professor, Yale, 1955–1985.
3. For a more lengthy discussion of the origins of the impossibility theorem, please see the following two sources: Kenneth Arrow, Amartya Sen, and Kotaro Suzumura, "Kenneth Arrow on Social Choice Theory," in *Handbook of Social Choice and Welfare*, vol. 2 (Elsevier BV, 2011), pp. 3–27; and Kenneth Arrow, "The Origins of the Impossibility Theorem," in *History of Mathematical Programming*, eds. Jan Karel Lenstra, Alexander H. G. Rinnooy Kan, and Alexander Schrijver (Amsterdam: Elsevier Science, 1991), pp. 1–5; also reprinted in this volume.
4. Michel Balinski and Rida Laraki, "A Theory of Measuring, Electing, and Ranking," *Proceedings of the National Academy of Sciences* 104, 21 (2007): 8720–8725.
5. S. S. Stevens, "On the Brightness of Lights and the Loudness of Sounds," *Science* 118 (1953): 576.
6. Daniel Kahneman and Alan Krueger, "Developments in the Measurement of Subjective Well-being," *Journal of Economic Perspectives* 20, 1 (2006): 3–24.
7. I elaborated on this point in my talk "Intergenerational Equity and the Rate of Discount in Long-term Social Investment," IEA World Congress, December 1995.

PART II

SUPPLEMENTAL MATERIALS

THE INFORMATIONAL BASIS OF SOCIAL CHOICE[1]

AMARTYA SEN

ABSTRACT

Any procedure of social choice makes use of some types of information and ignores others. For example, the method of majority decision concentrates on people's votes, but pays no direct attention to, say, their social standings, or their prosperity or penury, or even the intensities of their preferences. The differences between distinct procedures lie, to a substantial extent, on the kind of information that each procedure uses and what it has to ignore. The informational bases of the different social choice procedures tell us a great deal about how they respectively work and what they can or cannot achieve.

INTRODUCTION

Social choice theory addresses a wide range of decisional and judgmental problems, dealing with a variety of procedures—from voting to making normative social assessments. It encompasses theories of elections and balloting on one side, to welfare economics on the other, as well as theories of normative measurement, such as the evaluation of national incomes, measurement of inequality and poverty, and appraisal of social welfare. These distinct problems often demand quite dissimilar approaches, and there is little hope of getting some uniform approach that would work equally satisfactorily for all the different exercises.[2] Nevertheless, all the social choice problems have the shared feature of relating "social"—or group—assessment to the values, preferences, choices, or some other characteristics of the respective individuals who form the collectivity of that society or group. It is with the informational basis of that generic approach that this chapter is concerned.

Any procedure of social choice makes use of some types of information and ignores others. For example, the method of majority decision concentrates on people's votes and, by implication (given suitable assumptions), on their preferences. But it pays no direct attention to a variety of other kinds of information, such as their personal characters, their social standing, their prosperity or penury, or even the intensities of their preferences. But each of these ignored features can be very influential for other procedures of social choice. The differences between distinct procedures lie, to a substantial extent, on the kind of information that they respectively use and what they ignore. The informational bases of the respective social choice procedures tell us a great deal about their nature and underlying motivations, and about the differences that distinguish them from each other.

Much the same thing can be said about *principles* of social judgment or of normative social decisions. Their respective informational bases qualify the nature of the judgments that can—or cannot—be made. The informational bases can be explicitly stated, or implicitly characterized, by the axiomatic requirements imposed on each respective approach. Indeed, many of the "conditions" of social decisions that have received intense attention in recent decades, especially since Arrow's (1950, 1951) pioneering departure in axiomatic social choice theory, can be helpfully analyzed in terms of their informational demands, including informational prohibitions.

We can, in fact, go a long way in characterizing any normative principle, or any well-defined procedure of social choice, by identifying the information that it uses and, no less importantly, the information it ignores, as is discussed in Sen (1977b, 1979). It is tempting to ask whether a procedure, or a principle, of social choice can be fully delineated by its informational basis *alone*? The answer is very definitely in the negative, for reasons that are not far to seek. No matter how tight the informational conditions are that help us to move closely towards a certain principle *p*, there are at least three other principles that are not excluded by the very same informational requirements. First, a principle, let us call it anti-*p*, can demand the exact opposite of what principle concludes (e.g., minimizing the utility sum-total does not demand any more information than the utilitarian maximization of that sum). Second, to resolve that all alternatives be accepted as being equally good (what can be called *universal indifference*) does not demand any information at all (beyond the knowledge of that *blanket rule*). Third, resolving that no alternative can be ranked against any other (call it *universal unconnectedness*) also has little informational requirement.

However, the informational perspective can be used not only to make sure that there is enough information for the decisions involved, but also to demand that the processes or principles must actually respond to—or be sensitive to—particular types of information that characterize each specific principle or procedure. If we go beyond checking the sufficiency of available information to the necessity of taking discriminating note of particular information (a requirement of informational *responsiveness*), the last two of the three alternatives may turn out to be unacceptable. To illustrate with a simple example, an insistence on the necessity to be sensitive to the unanimous strict preferences of all individuals over a pair of social states $\{x, y\}$ may eliminate the possibility of being indifferent between them, or keeping them unranked, when everyone does in fact prefer x to y. But informational sensitivity alone, without a directional mandate, cannot yield the conclusion that x must, in this case, be socially preferred to y (or that x must be chosen over y, given that choice). Indeed, the exact opposite (that is, socially ranking y over x in response to every member of the society preferring x to y), while patently perverse, is also sensitive to the same information. Pareto and anti-Pareto do not, in fact, differ in informational requirement or in informational responsiveness—only in the *direction* of the response.

Substantive social choice theory cannot, therefore, be understood merely in terms of informational bases. There is no way of getting rid of the different ways in which the same information can be used in different procedures or principles. And yet the differences between distinct social choice principles or processes often rest primarily on their contrasting informational bases.

Indeed, many of the decisive steps in the history of social choice theory have turned on invoking some types of information and disavowing others. This chapter is concerned with tracing some of these historical steps, beginning with the early origins of the subject in the hands of eighteenth-century French mathematicians (such as Condorcet and Borda), then proceeding to the founding of modern social choice theory in the middle of the twentieth century (with the pioneering work of Kenneth Arrow in 1950 and 1951), and finally looking briefly at some of the more recent trends in the subject, particularly in normative social choice theory.[3]

DEMOCRATIC PRIORITIES AND INFORMATIONAL INCLUSIVENESS

The informational foundation of modern social choice theory relates closely to the basic democratic conviction that social judgments and public decisions must depend, in some transparent way, on individual preferences, broadly understood. The emergence of this democratic instinct relates closely to the ideas and events that surrounded the European Enlightenment. While it drew on various antecedent sources and inspirations, the democratic perspective received adequate delineation and wide public acknowledgement only during the Enlightenment, particularly in the late-eighteenth century, which also saw the French Revolution and American independence. The works of Borda (1781) and Condorcet (1785) on the properties of voting systems and of Bentham (1789) on the demands of utilitarian social aggregation were clearly influenced by this general intellectual climate.

Individual preferences can, of course, be variously interpreted in different democratic exercises, and this is well illustrated by the contrast between focusing on votes—and through them on people's voices (explored in the classic works of Borda or Condorcet)—on the one hand, and concentrating on the interests or judgments (explored in the pioneering writings of Adam Smith, Jeremy Bentham, or John Stuart Mill), on the other. These contrasts—between alternative interpretations of preferences—can be very important for some purposes, but in the present context, that of noting the democratic foundations of social choice theory, their differences are less important than what they have in common. For the moment I shall use the generic term *preference* to cover all these different interpretations of individual concerns that could be invoked, in one way or another, to serve as the informational bases of public decisions or social judgments.

In this elementary democratic foundation, there is also a strong understanding that at least in principle the preferences of everyone could count, without any a priori exclusions (even if it were to turn out, on the basis of further analyses, that some preferences would be more effective than others, sometimes radically so). No member of the collectivity could be eliminated as being foundationally irrelevant. That informational underpinning, which was established in the first generation of formal social choice theory that emerged in the eighteenth century, has been powerfully reinforced in modern social choice theory pioneered by Arrow (1951). While this rudimentary democratic feature may seem today to be rather straightforward, perhaps even mundane, it was a radical enough step at the time of its first consolidation. It firmly established a democratic inclusiveness in the informational basis of social choice.

The force of this departure can be brought out by examining the contrast between these new presumptions and the earlier interpretations of social choice in pre-Enlightenment politics or political economy. For example, Aristotle in ancient Greece and Kautilya in ancient India, both of whom lived in the fourth century before Christ, had explored various procedures for making appropriate social decisions, in their books respectively called *Politics* and *Economics*.[4] Aristotle had no great difficulty in excluding women and slaves from the process of social decision making, and Kautilya had little problem in confining critical decisions in the hands of those blessed by high social status (related to caste or political authority). Enormous informational exclusions were, in effect, authorized *before* the substantive investigation of social procedures properly began. The world of late-eighteenth century Europe broadened the needed informational foundation by overturning that exclusionary authorization.

The pioneers of modern social choice theory were guided by their firm conviction that every member of a collectivity must, in principle, count in the decisions of that collectivity.[5]

This is meant to apply to choices for the society as a whole, as well as to those for a committee, or an academy, or a set of juries, or some other organizational panel or board. There is an insistence on no antecedent exclusion of the concerns of any individual member of the respective collectivity. The different decisional structures that were explored, such as majority rule, or utilitarian aggregation, or rank-order voting (the so-called "Borda rule") differed in the significance that can be attached to particular aspects of individual preferences: for example, whether to take note only of the ranking of each pair considered separately (ignoring other "irrelevant" alternatives), or to attach significance to the rank of a particular alternative in an overall ranking (as in Borda's well-known formula).[6] But in not excluding anyone's preference ordering from counting, they shared a basic belief in democratic inclusiveness in an elementary but powerful form.

Since contemporary social choice theory, pioneered by Arrow, emphatically shares this foundational democratic value, the discipline has continued to be loyal to this basic informational presumption. For example, when an axiomatic structure yields the existence of a dictator (Arrow 1951), as a joint implication of chosen axioms that seemed plausible enough (seen on their own), this is immediately understood as something of a major embarrassment for that set of axioms, rather than being taken to be just fine on the ground that it is a logical corollary of axioms that have been already accepted and endorsed. We cannot begin to understand the intellectual challenge involved in Arrow's impossibility theorem without coming to grips with the focus on inclusiveness that goes with a democratic commitment, which is deeply offended by a dictatorial procedure, even when it is entailed by axiomatic requirements that seem eminently acceptable.

The same applies, in one way or another, to the various subsequent results that followed Arrow's impossibility theorem. Arrow's specific impossibility result—with dictatorship's being implied by his other conditions—is not extendable to the case in which the transitivity of social preference is weakened, even rather slightly, to just "quasitransitivity" (or the transitivity of strict preference only), unless the other conditions are redefined.[7] But as the requirement of "collective rationality" in the form of social transitivity or binariness is gradually relaxed, new results emerge that show that there must now be an oligarchy, or the existence of someone with veto power, or some other violation of what democracy demands.[8] Again, the tension that is generated by these results relates to the violation of informational inclusion that a democratic commitment entails.

It is important to recognize how radically the nature of the social choice search for a minimally acceptable social decision procedure has been shaped by the informational implications of a basic democratic conviction that was getting firmly established in the second half of the eighteenth century, just as social choice theory, in its early form, was being founded. Given the centrality of Arrow's result, which has profoundly shaped the direction that the development of social choice theory took over the second half of the twentieth century, it is perhaps useful to examine sequentially how the tension with democratic inclusiveness emerges from what looks merely like minimal demands of systematic and sensitive social choice. After some clarificatory discussion of the informational aspects of Arrow's social choice framework in the next section, a simple way of understanding and establishing Arrow's theorem (seen specifically in an informational perspective) is presented in section 4.

INFORMATIONAL EXCLUSIONS AND SOCIAL CHOICE FRAMEWORK

In the general Arrovian framework, the social ranking R of the alternative social states is taken to be a function of the n-tuple of individual rankings $\{R_i\}$ of those states:

$$R = f(\{R_i\}) \tag{1}$$

The functional relation f, which we can call a "collective choice rule," is an Arrovian "social welfare function" when there is the further requirement that R as well as each R_i be a complete ordering of feasible social states.[9] In the discussion that follows, the immediate reference will be specifically to social welfare functions (SWF), but much of the discussion applies to collective choice rules in general.

Since conflict with informational inclusiveness is such a central feature of the Arrow impossibility theorem, it is important to be sure that the impossibility result is not being achieved simply by beginning with patently informational restrictions (in the formulation of social welfare functions). We have to ask: How inclusive is this general Arrovian formulation regarding the information that can be accommodated in the process of social choice? In answering this question, it is convenient to distinguish between "utility information" in the general sense (including information about preference rankings) and "nonutility information" regarding other features of states of affairs. It is easily checked that while the utility information that is allowed to be accommodated in an SWF is rather restricted, there is nothing in the form of an SWF itself that limits the admissibility of nonutility information.

Consider first the nature of allowable utility information. In (1), the form of $f(\{R_i\})$ does not allow the use of interpersonal comparison of utilities.[10] This is certainly a start-off restriction. In his initial formulation of the problem of social choice, Kenneth Arrow was moved by the view, common in positivist philosophy that was then influential in welfare economics, that "interpersonal comparison of utilities has no meaning" (Arrow 1951, p. 9). The utility information that is usable in this structure of social choice consists of n-tuples of individual preferences (or utility orderings) of the respective individuals—considered separately. This is a momentous informational exclusion, the removal of which can open up many constructive possibilities (as was discussed in Sen 1970a).

However, as far as nonutility information is concerned, the format of social welfare functions is remarkably permissive. Unless eliminated by specific axioms to be imposed on social welfare functions (on which more will be discussed presently), the framework can accommodate sensitivity to any part of the informational content of social states. There is a real comprehensiveness here, which is worth emphasizing, since it can be easily missed because of the apparent insistence, in the formulation of (1), that the *n*-tuple of individual preferences $\{R_i\}$ be the sole input into the choice process. The implications of this formulation require some elucidation, particularly since they are, in fact, critically important for later social choice theory, involving the use of nonutility information related to liberties, rights, and nonwelfarist interpretations of justice, equity, and poverty.

The informational content of social states is not arbitrarily restricted in any way, and the social welfare function can take note of any information that can be accommodated within the specification of social states. There is need for some clarification here. Given the form of (1), with $R = f(\{R_i\})$, it may appear that no feature of social states can be influential in the choice over these states unless the individual preferences $\{R_i\}$ respond to that feature. In this interpretation, if a specific feature of a state of affairs (say, the level of income inequality, or the violation of some liberties, or the infringement of civil rights) is going to be directly influential in social choice, it must be *through* the impact of that feature on individual preferences over states of affairs. Indeed, in this interpretation, no feature of the states can have an influence on social choice through any channel *other than* individual preferences.

This interpretation, however, is not correct. Even though a social welfare function insists on a tight functional relation f between R and $\{R_i\}$, there is nothing in the mathematics of this requirement that would prevent the nature of the functional relation f to be itself responsive to any information that is included in the content of the respective social states that R and R_i order. The individual preference orderings are rankings of substantive social states, and the information about the social states can be taken into account in deciding on the mapping between the set of n-tuples of individual preferences and the set of social rankings to be determined by f.

However, this permissive format can be made informationally more restrictive through the effects of axioms that may be imposed on a social welfare function. Indeed, through this route it is possible to end up eliminating the direct usability of *all* nonutility information, so that the characteristics of social states are made totally inconsequential: social choice over them will then be determined only by their placing in the individual preference rankings. This condition is sometimes called "neutrality." This is perhaps an oddly reverent name for what is after all only an informational restriction, but the requirement can be seen as "neutralizing"—indeed eliminating—the influence of all nonutility (or nonpreference) information regarding social states. In effect, it yields an insistence that social decisions be taken only on the basis of individual preferences over the states, without paying any attention to the nature of these states (and the nonutility or nonpreference information about these states).

In the literature of moral philosophy, this "neutrality" condition and similar requirements are sometimes called "welfarism" (see Sen and Williams 1982), and that term has been in increasing use in social choice theory as well. Welfarism, narrowly defined, is the demand that social welfare (or whatever is taken as the social maximand) depends only on individual utilities: Other features of states of affairs have no direct influence on social welfare (or the social maximand). In somewhat broader formulations, welfarism, corresponding to "neutrality," can be seen as a more permissive insistence that the social maximand depends only on individual utilities, or individual welfares, or individual evaluations of the worth of states of affairs (more on this presently).

It is possible to combine welfarism with very rich utility information (such as interpersonal comparability and cardinality), and indeed such enrichment of information would be particularly important for normative social judgments, including welfare economic assessments.[11] But when it is applied to social welfare functions that use individual utility information only in the form—as in (1)—of *n*-tuples of individual preferences (corresponding to noncomparable ordinal individual utilities), we get a combination that attempts to make do with very little information indeed. It must, however, be noted that Arrow does not invoke neutrality or welfarism in any form as a prior requirement. In a limited form that restriction emerges as an implication of other conditions, and a substantial part of the unexpected nature of Arrow's impossibility result (or the "General Possibility Theorem" as Arrow called it) relates, in fact, to this analytical demonstration.[12]

AXIOMATIC EXCLUSIONS AND ARROW'S IMPOSSIBILITY THEOREM

ARROW'S THEOREM

It is useful, in this perspective, to go through a simple proof of Arrow's impossibility theorem not merely because the result is so central to social choice theory, but also because of the light it throws on the way apparently mild axioms can, acting in combination, end up as very severe informational constraints. Arrow considered a set of very plausible-looking conditions relating social choice to the n-tuple of individual preferences, and showed that it is impossible to satisfy those conditions simultaneously.

The axioms used by Arrow (in the later, and neater, version in Arrow 1963) include *unrestricted domain, weak Pareto Principle, nondictatorship,* and *the independence of irrelevant alternatives* (in addition to the structural conditions requiring that the set of individuals is finite and that the set of social states includes at least three distinct states). We define xR_iy as the statement that person i weakly prefers x to y (that is, either strictly prefers x to y, or is indifferent between them), and xP_iy as person i strictly prefers x to y. The weak and strict social preferences are denoted R and P, respectively.

Unrestricted domain (U) demands that the domain of the social welfare function, that is f in (1), includes all possible n-tuples of individual preferences $\{R_i\}$. The *weak Pareto Principle* (P) says that if all persons prefer any x to any y, then x is socially preferred to y. *Nondictatorship* (D) excludes the possibility that any individual j could be so powerful that whenever, over the domain of f, he or she prefers any x to any y, society too strictly prefers x to y. And *independence of irrelevant alternatives* (I) can be seen as demanding that the social ranking of any pair $\{x, y\}$ must depend only on individual preferences over $\{x, y\}$. The Arrow impossibility theorem states that there does not exist any social welfare function f that can simultaneously fulfill U, P, D, and I.

We can define a set G of individuals as being "decisive" over the ordered pair $\{x, y\}$, denoted $DG(x, y)$, if and only if whenever everyone in G prefers x to y, we must have xPy no matter what others prefer. Further, if G is decisive over every ordered pair, then G is simply called "decisive," denoted DG. It is readily seen that nondictatorship is the requirement that no individual is decisive, whereas the weak Pareto Principle is the requirement that the set of all individuals is decisive. The proof used here goes via two lemmas, which establish the implied informational exclusions, to obtain dictatorship from the weak Pareto Principle.[13]

PROOF OF ARROW'S THEOREM

Lemma L.1 *If $DG(x, y)$ for any ordered pair $\{x, y\}$, then DG.*

To establish this, we have to show that $DG(x, y) \rightarrow DG(a, b)$, for all a and b. The demonstration proceeds by repetitions of essentially the same strategy in different possible cases depending on whether or not x or y is identical with either b or a. Consider the case in which the four states x, y, a, b are all distinct. Assume the following pattern of individual preferences: for all persons j in G: $aP_j x$, $xP_j y$, and $yP_j b$, and for all persons i not in G: $aP_i x$ and $yP_i b$

(with nothing being presumed about the ranking of the other pairs). By $DG(x, y)$, we have xPy, and by the weak Pareto Principle, we obtain aPx and yPb. Hence, through the transitivity of strict preference, we get: aPb. By the independence of irrelevant alternatives, aPb must depend on individual preferences only over $\{a, b\}$, and since only the preferences of people in G have been specified, clearly $DG(a, b)$.

Lemma L.2 *For any G, if DG, and if G has more than one person in it and can be, thus, partitioned into two nonempty parts* G_1 *and* G_2*, then either* DG_1 *or* DG_2.

Assume that for all i in G_1: xP_iy, and xP_iz, with any possible ranking of y, z, and that for all j in G_2: xP_jy, and zP_jy, with any possible ranking of x, z. Nothing is required from the preferences of those not in G. Clearly, xPy by the decisiveness of G. If, now, xPz, then group G_1 would be decisive over this pair, since they alone definitely prefer x and z (the others can rank this pair in any way). If G_1 is not to be decisive (and thus by Lemma L.1 not to be decisive over any pair), we must have zRx for *some* set of individual preferences over x, z of nonmembers of G_1. Take that case. So we have zRx and also xPy. We thus have, by transitivity of preferences, zPy. Since only G_2 members definitely prefer z to y, this entails that G_2 is decisive over this pair $\{z, y\}$. But, then, by L.1, G_2 is generally decisive. So either G_1 or G_2 must be decisive.

Completing the proof of Arrow's theorem can now proceed very rapidly. By the weak Pareto Principle, the group of all individuals is decisive. It is, by assumption, finite. By successive twofold partitionings, and each time picking the decisive part (which exists, guaranteed by L.2), we arrive at a decisive individual, who must, thus, be a dictator.

INTERPRETATION OF THE PROOF

The proof just presented works through a sequential compounding of informational exclusions, beginning in a small way and ending with such a massive prohibition that the weak Pareto Principle cannot be effectively distinguished from the existence of a dictator. In socially ranking *x* and *y*, the condition of independence of irrelevant alternatives excludes the use of information—both preferences related and any other—except what relates directly to the "relevant" alternatives, that is, *x* and *y* only. Starting with that small beginning and afforced by the unrestricted domain and the weak Pareto Principle, we get to the result in Lemma L.1 that any group that is socially decisive over any pair of social states must be socially decisive over every pair of social states—no matter what these states are. So the specific information about the respective states, which we may have plentifully, will not be allowed to make any difference as far as decisiveness (based on individual preferences) is concerned.

Armed with this informational exclusion established in Lemma L.1, the proof proceeds in Lemma L.2 to economize also on preference information itself. If the information about the unanimous preference of a set *G* of individuals is adequate to rank social states (no matter what others want), then the information about the unanimous preference of some proper subset G^* of individuals, excluding some others in *G*, will be adequate as well. This opens the door, through sequential use, to go from the decisiveness of unanimous strict preference of all (thanks to the weak Pareto Principle) to the decisiveness of strict preference of some one individual (that is, a dictator). The informational inclusiveness of a foundational democratic commitment is, thus, caught in a fierce internal contradiction: to empower *all* without discrimination (as incorporated in the weak Pareto Principle) is to empower *one* irrespective of what others want (that is, a dictatorship).

ENRICHING INFORMATION FOR THE POSSIBILITY OF SOCIAL CHOICE

As was mentioned earlier, the *entailed* exclusion of nonutility information through Arrow's axioms adds to the *antecedent* exclusions directly incorporated in the formulation of a social welfare function through the nonadmissibility of interpersonal comparability and cardinality. Informational enrichment can be sought *either* through the route of enriching utility information *or* through that of admitting nonutility information.

The former route has been particularly explored in social choice theory since the 1970s.[14] The formal structure is that of a "social welfare functional," which functionally relates social ranking R to n-tuples of individual utility functions $\{Ui\}$. The extent of measurability and cardinality of utilities is specified by "invariance conditions" imposed on social welfare functionals.[15] It can be shown that all the Arrow axioms, if translated into a broadened framework of "social welfare functionals" that allows richer utility information, can be simultaneously satisfied, even with just ordinal comparability of individual utilities, even without any cardinality.

Furthermore, many other constructive possibilities are opened up once cardinal comparability is also allowed.[16] These extensions are hugely important for welfare economics and for normative social judgments in general.

This route (that is, enriching utility information while still keeping out nonutility information) does not immediately raise issues of inconsistency. But while the Arrow impossibility may be circumvented this way, this path does not still accommodate the use of nonutility information needed for specification of rights or liberties or nonwelfarist assessment of inequality or fairness. This is unfortunate since these norms have considerable appeal, and many "non-neutral" concerns have figured in the informal literature on social decisions and choices for a very long time. Even as formal social choice theory was getting founded through the pioneering works of Borda (1781), Condorcet (1785), and other mathematical analysts of voting and electoral processes, and also through the parallel line of investigation pursued by utilitarians such as Bentham (1789), other innovative departures were being made in the understanding of justice in a way that could not be fully translated into axioms defined in the neutral space of preferences. For example, the relevance and reach of the idea of rights were extensively explored by such pioneering authors as Mary Wollstonecraft (1790, 1792) and Thomas Paine (1791).[17] These concentrations were well reflected in the practical politics related to the French Revolution as well as American Independence, both of which made extensive use of the idea of fundamental rights.

Some of these rights involve conditions that relate to individual preferences of the people involved (for example, what one prefers in one's "personal domain" of liberty), but even here nonutility characteristics of the states of affairs have to be taken into account to give more effectiveness to each person over his or her own personal domain. These concerns were not explicitly accommodated in formal social choice theory in its classic formulations, but they have figured prominently in more recent developments in the social choice literature.[18]

ON COMBINING UTILITY AND NONUTILITY INFORMATION

There are, however, important problems in combining utility and nonutility information, since their disparate roles can yield possible inconsistencies. Indeed, this is one way of interpreting the so-called liberal paradox, which involves a consistency problem in simultaneously accommodating, along with unrestricted domain, a minimal condition of liberty (involving the use of some nonutility information regarding personal features in social states) and the weak Pareto Principle (involving very modest use of utility information).[19] Indeed, the "impossibility of the Paretian liberal" brings out this tension in a very simple case, and the conflicts can be much more complex when it is attempted to use richer utility information *along with* substantive use of nonutility descriptions of states of affairs.

These "hybrid" frameworks have not been extensively investigated yet, and there has in fact been some reluctance to leave the simplicity of welfarism even when trying to accommodate principles or procedures of social choice that are quintessentially nonwelfarist. Consider, for example, John Rawls's (1971) well-known theory of justice, which involves "the priority of liberty" as the first principle (a substantively nonwelfarist requirement), and also the Difference Principle, which uses lexicographic maximin in the space of primary goods, not utilities. While Rawls has been much invoked in social choice theory, nevertheless the axiomatizations of Rawls in welfare economics (and in social choice theory related to welfare economics) have tended to ignore his first principle altogether (except indirectly in the context of the so-called liberal paradox) and have also redefined the Difference Principle in terms of utilities, in contrast with Rawls's own focus on primary goods.[20]

These recharacterizations of nonwelfarist principles (like Rawls's) in "welfarist" terms, while strictly speaking inaccurate, do have significant usefulness, for several distinct reasons. First, welfarism does appeal to the intuition of many social analysts.[21] In fact, many seem to find the welfarist version of Rawls's lexicographic maximin more acceptable than Rawls's own insistence on operating in the space of primary goods.[22]

Second, the utility-based formulation is open to alternative interpretations and can be relatively easily integrated with decision-theoretic normative reasoning. Indeed, as d'Aspremont and Gevers (2002) point out in their masterly critical survey of the literature on "social welfare functionals" in the first volume of the *Handbook of Social Choice and Welfare* this literature "can be reinterpreted as an application of multi-objective decision theory to the ethical observer's problem" (p. 464).

Third, the restricting of the entire informational basis of all the normative principles to one basic class of evaluative data (such as individual utilities or individual overall evaluations) can be a simple way of keeping the possibility of inconsistency (under discussion here) at bay, and this can be seen to be a considerable merit in itself. Indeed, even within such restrictions, a variety of different concerns can be accommodated, without internal tension. As d'Aspremont and Gevers (2002) point out, even the alternative approach of capabilities (on which see chapter 16 by Kaushik Basu and Luis Felipe Lopez-Calva in *Handbook of Social Choice and Welfare*, vol. 1) can do this with its exploration of the possibility of accommodating "doings and beings" *within* the approach of "social welfare functionals." The problems that may have to be faced would arise not from the *objects* that influence individual utility or individual welfare or individual evaluation, but from any proposed use of other information to determine the relevance of—and the weights to be placed on—utility or welfare, or on evaluative conclusions of individuals.

Possible tensions arise when *other* data from the states of affairs are invoked in making social judgments or social choices: for example, in giving priority to a person's evaluation over her own "personal domain" (as in Rawlsian "priority of liberty" or in various conditions of "minimal liberty" used in social choice theory), or in attaching special importance to the centrality of certain capabilities (as in the philosophical or developmental approaches that give a special role to the fulfilment of certain basic capabilities).[23] The critical issue, as was discussed earlier, is so-called neutrality—a condition that can be directly imposed or indirectly precipitated through the use of other axioms, which restricts the class of permissible social welfare functions and social welfare functionals to reliance on utility information.[24] This requirement runs counter to the invocation of various foundational norms, reflected in, say, the Rawlsian principles of justice, or the Aristotelian focus on capabilities, or the Wollstonecraft–Paine concentration on the "vindication of rights."

The possibility of combining these different classes of foundational information reflected in different types of principles of social justice and equity—involving both utility and nonutility information—has not yet been investigated adequately. While celebrating what has already been achieved, it is important to identify this as an area in which more investigation will be needed in the future.[25] That would be of particular importance in making further use of social choice theory in analyzing and exploring theories of justice.

NOTES

1. For helpful discussions I am grateful to Kenneth Arrow and Kotaro Suzumura.
2. I have discussed elsewhere, in Sen (1977a), why the distinctions involved are very important to acknowledge and accommodate. See also Suzumura (1982, 1983).
3. The balance of attention in this chapter is somewhat tilted in the direction of relatively broader results rather than more particularized findings. I have had the occasion to discuss and illustrate in another essay on social choice theory (Sen, 1986) how the classical debates on the principles and procedures of social choice have continuing relevance to the formal analyses and

technical results in the more recent works that often deal with very specialized issues. The "older" contentious issues, related to informational bases, have similar resilience and durability.

4. For English translations of Aristotle's *Politics* and of Kautilya's *Arthashastra*, see, respectively, Barker (1958) and Shama Sastry (1967). *Arthashastra*, the Sanskrit title of Kautilya's book, which literally means "treatise on material wealth," is perhaps best translated as "economics," even though much of the book is devoted to studying systematic statecraft.
5. On the intellectual debates that engaged Enlightenment authors, including Condorcet, see Rothschild (2001).
6. The so-called "Borda rule" belongs to a general class of "positional rules," the properties of which have been extensively investigated by Gardenfors (1973), Fine and Fine (1974), and others; see Pattanaik's (2002) superb critical survey of this literature, in Vol. 1 of *Handbook of Social Choice and Welfare*, vol. 1. Condorcet's voting principles have been well discussed by Arrow (1963), Fishburn (1973), Suzumura (1983), and Young (1988), among others.
7. This nonextendability is shown in Sen (1969, 1970a). However, it can also be shown that in a choice-functional framework, Arrow's conditions can be suitably recharacterized, without changing their motivational justifications, to precipitate the dictatorship result without any condition of internal consistency of social choice whatsoever (as shown in Sen 1993). Thus reinterpreted, Arrow's conditions of independence, unrestricted domain, and the weak Pareto Principle, together, continue to contradict democratic inclusiveness, even without any demand for *internal* consistency of social choice.
8. Different types of results in this general line of investigation have been presented—or scrutinized—in Gibbard (1969, 1973), Sen (1970a, 1977a, 1993), Mas-Colell and Sonnenschein (1972), Fishburn (1973, 1974), Brown (1974, 1975), Binmore (1975, 1994), Campbell (1976), Deb (1976, 1977), Suzumura (1976a, b, 1983),

Blau and Deb (1977), Kelly (1978), Blair and Pollak (1979, 1982), Grether and Plott (1982), Chichilnisky (1982), Chichilnisky and Heal (1983), Moulin (1983), Pattanaik and Salles (1983), Peleg (1984), Hammond (1985), Kelsey (1985), and Campbell and Kelly (1997).

9. Different types of collective choice rules can be distinguished and individually investigated, as discussed in Sen (1970a). See also Fishburn (1973), Kelly (1978), and Suzumura (1983), among other treatises, for different classificatory systems.

10. Nor are cardinal utilities (whether or not comparable) admissible in this framework. This is not, however, a critical constraint for Arrow's impossibility theorem, since that result can be extended to the case of noncomparable cardinal utilities (see Theorem 8.2 in Sen 1970a), even though it does severely restrict the class of permissible social choice procedures (on which see Gevers 1979 and Roberts 1980a).

11. This general issue is discussed more fully in Sen (1970a, 1977a, 1999).

12. Since it is often said that Arrow's impossibility theorem is a generalization of the old "paradox of voting," it is worth noting that this is only partly true. It is important, in particular, to recognize that while voting rules must satisfy neutrality, neither the form of Arrow's social welfare function, nor any of the individual axioms imposed by Arrow on that function, make any demand of neutrality. We are moved in the direction of neutrality by the combination of the axioms, and indeed the main work in proving Arrow's impossibility theorem consists, it can be argued, in deriving a property of neutrality from the *combination* of different axioms, through the use of axiomatic reasoning. In what follows, see Lemma L.1.

13. The brief proof used here corresponds to the one outlined in Sen (1995, fns. 9 and 10, p. 4). Note that this proof drops the need to introduce the intermediate concept of "almost decisiveness" (used in Arrow 1951, 1963, and Sen 1970a), since it is redundant.

14. Vickrey (1945) and Harsanyi (1955) had earlier identified representation results that make use of information based on expected utility, interpreted as interpersonally comparable cardinal utilities, to obtain the value of social welfare in a summational form.
15. The operation and use of invariance conditions is investigated in Sen (1970a, b), d'Aspremont and Gevers (1977), Gevers (1979), Maskin (1979), and Roberts (1980a), among other contributions.
16. See, among other contributions, Sen (1970a, 1977b), Hammond (1976, 1985), d'Aspremont and Gevers (1977, 2002), Arrow (1977), Maskin (1978, 1979), Gevers (1979), Roberts (1980a, b), Suzumura (1983, 1996), Blackorby, Donaldson, and Weymark (1984), d'Aspremont (1985), Blackorby, Bossert, and Donaldson (2002), and d'Aspremont and Mongin (2008).
17. The far-reaching relevance of these perspectives is discussed in Sen (2009).
18. See Arrow, Sen, and Suzumura (1996/1997) and other chapters of *Handbook of Social Choice Theory*, vol. 2, especially part 7, including the chapters by William Thomson, Marc Fleurbaey and Francois Maniquet, and Kotaro Suzumura.
19. The literature on the liberal paradox is by now quite vast. The special number of *Analyse & Kritik*, September 1996, includes a fine collection of papers on this subject, as well as extensive bibliographies of publications in this field
20. See Phelps (1973, 1977), Hammond (1976), Maskin (1978, 1979), Meade (1976), Strasnick (1976), Arrow (1977), d'Aspremont and Gevers (1977), Sen (1977b), Gevers (1979), Roberts (1980a, 1980b), Atkinson (1983), Suzumura (1983, 1996), Blackorby, Donaldson, and Weymark (1984), d'Aspremont (1985). However, social choice theoretic reasoning has been used to question or defend the possibility of having a consistent index of bundles of diverse primary goods, on which see Plott (1978), Gibbard (1979), Blair (1988), and Sen (1991).

21. On this see Hammond (1976, 1982), d'Aspremont and Gevers (1977, 2002), Gevers (1979), Suzumura (1983, 1996), Broome (1991, 2004), Pattanaik and Suzumura (1994), d'Aspremont and Mongin (2008), Blackorby, Bossert, and Donaldson (2002), among other contributions.
22. I should explain that this is definitely not the judgment of this author, but he is able to distinguish between his own assessment and that of many analysts whose judgments he respects.
23. See Sen (1982), Atkinson (1983), Suzumura (1983), Nussbaum and Sen (1993), Dutta (2002), among other contributions.
24. In the format of social welfare functionals, strong neutrality follows from unrestricted domain, Pareto indifference, and independence in a binary form, as was established by d'Aspremont and Gevers (1977); see Theorem 3.7 in d'Aspremont and Gevers (2002, pp. 493–494). The conditions that yield something quite close to neutrality for social welfare functions, as shown in Lemma L.1 above, draw on similar informational demands imposed on that framework.
25. This has been discussed in the exploration of social choice–based ideas of justice in Sen (2009).

REFERENCES

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, *58*.

Arrow, K. J. (1951). *Social choice and individual values*. New York: Wiley.

Arrow, K. J. (1963). *Social choice and individual values*, (enlarged 2nd ed.). New York: Wiley.

Arrow, K. J. (1977). Extended sympathy and the possibility of social choice. *American Economic Review*, *67*.

Arrow, K. J., & Intriligator, M. (Eds.). (1986). *Handbook of mathematical economics*. Amsterdam: North-Holland.

Arrow, K. J., Sen, A. K., & Suzumura, K. (1996/1997). *Social choice re-examined*. London: McMillan, and New York: St. Martin's Press.

Atkinson, A. B. (1983). *Social justice and public policy*. Brighton: Wheatsheaf, and Cambridge, MA: MIT Press.

Barker, E. (1958). *The politics of Aristotle*. London: Oxford University Press.

Bentham, J. (1789). *An introduction to the principles of morals and legislation*. London: Payne; republished, Oxford: Clarendon Press, 1907.

Binmore, K. (1975). An example in group preference. *Journal of Economic Theory*, *10*.

Binmore, K. (1994). *Playing fair: Game theory and the social contract* (Vol. 1). Cambridge, MA: MIT Press.

Blackorby, C., Donaldson, D., & Weymark, J. (1984). Social choice with interpersonal utility comparisons: A diagrammatic introduction. *International Economic Review*, *25*.

Blackorby, C., Bossert, W., & Donaldson, D. (2002). Utilitarianism and the theory of justice. In K. J. Arrow, A. K. Sen & K. Suzumura (Eds.), *Handbook of social choice and welfare* (Vol. 1). Amsterdam: North-Holland.

Blair, D. H. (1988). The primary-goods indexation problem in Rawls's theory of justice. *Theory and Decision*, *24*.

Blair, D. H., & Pollak, R. A. (1979). Collective rationality and dictatorship: the scope of the Arrow theorem. *Journal of Economic Theory*, *21*.

Blair, D. H., & Pollak, R. A. (1982). Acyclic collective choice rules. *Econometrica*, *50*.

Blau, J. H., & Deb, R. (1977). Social decision functions and veto. *Econometrica*, *45*.

Borda, J. C. (1781). Mémoire sur les elections au Scrutin, *Histoire de l'Académie Royale des Sciences* (Paris); translated by Alfred de Grazia, Mathematical Derivation of an Election System, *Isis*, *44*.

Broome, J. (1991). *Weighing goods: Equality, uncertainty and time*. Oxford: Blackwell Press.

Broome, J. (2004). *Weighing lives*. Oxford: Oxford University Press.

Brown, D. J. (1974). An approximate solution to Arrow's problem. *Journal of Economic Theory, 9*.

Brown, D. J. (1975). Acyclic aggregation over a finite set of alternatives, Cowles Foundation Discussion Paper No. 391, Yale University.

Campbell, D. E. (1976). Democratic preference functions. *Journal of Economic Theory, 12*.

Campbell, D. E., & Kelly, J. S. (1997). The possibility-impossibility boundary in social choice. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Social choice re-examined* (Vol. 1). New York: St. Martin's Press.

Chichilnisky, G. (1982). Topological equivalence of the Pareto condition and the existence of a dictator. *Journal of Mathematical Economics, 9*.

Chichilnisky, G., & Heal, G. (1983). Necessary and sufficient conditions for resolution of the social choice paradox. *Journal of Economic Theory, 31*.

de Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: L'Imprimerie Royale.

d'Aspremont, C. (1985). Axioms for social welfare ordering. In L. Hurwicz, D. Schmeidler, & H. Sonnenschein (Eds.), *Social goods and social organization*. Cambridge: Cambridge University Press.

d'Aspremont, C., & Gevers, L. (1977). Equity and the informational basis of collective choice. *Review of Economic Studies, 44*.

d'Aspremont, C., & Gevers, L. (2002). Social welfare functionals and interpersonal comparability. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare* (Vol. 1). Amsterdam: North-Holland.

d'Aspremont, C., & Mongin, P. (2008). A welfarist version of Harsanyi's aggregation theorem. In M. Fleurbaey, M. Salles, & J. A. Weymark (Eds.), *Justice, political liberalism and utilitarianism: Themes from Harsanyi and Rawls*. Cambridge: Cambridge University Press.

Deb, R. (1976). On constructing generalized voting paradoxes. *Review of Economic Studies*, *43*.

Deb, R. (1977). On Schwartz's rule. *Journal of Economic Theory*, *16*.

Dutta, B. (2002). Inequality, poverty and welfare. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *The handbook of social choice and welfare* (Vol. 1). Amsterdam: North-Holland.

Fine, B. J., & Fine, K. (1974). Social choice and individual ranking. *Review of Economic Studies*, *41*.

Fishburn, P. C. (1973). *The theory of social choice*. Princeton, NJ: Princeton University Press.

Fishburn, P. C. (1974). On collective rationality and a generalized impossibility theorem. *Review of Economic Studies*, *41*.

Gardenfors, P. (1973). Positional voting functions. *Theory and Decision*, *4*.

Gevers, L. (1979). On interpersonal comparability and social welfare orderings. *Econometrica*, *47*.

Gibbard, A. F. (1969). Unpublished term paper at the Philosophy Department of Harvard University, discussed in Sen (1970a).

Gibbard, A. F. (1973). Manipulation of voting schemes: a general result. *Econometrica*, *41*.

Gibbard, A. F. (1979). Disparate goods and Rawls' difference principle: A social choice theoretic treatment. *Theory and Decision*, *11*.

Grether, D. M., & Plott, C. R. (1982). Nonbinary social choice: An impossibility theorem. *Review of Economic Studies*, *49*.

Hammond, P. J. (1976). Equity, Arrow's conditions and Rawls' difference principle. *Econometrica*, *44*.

Hammond, P. J. (1982). Liberalism, independent rights, and the Pareto Principle. In L. J. Cohen, J. Los, H. Pfeiffer, & K.-P. Podewski (Eds.), *Logic, methodology, and the philosophy of science* (Vol. 6). Amsterdam: North-Holland.

Hammond, P. J. (1985). Welfare economics. In G. Feiwel (Ed.), *Issues in contemporary microeconomics and welfare*. Albany: SUNY Press.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics and interpersonal comparisons of utility. *Journal of Political Economy*, *63*.

Kelly, J. S. (1978). *Arrow impossibility theorems*. New York: Academic Press.

Kelsey, D. (1985). Acyclic choice without the Pareto Principle. *Review of Economic Studies*, *51*.

Mas-Colell, A., & Sonnenschein, H. (1972). General possibility theorems for group decisions. *Review of Economic Studies*, *39*.

Maskin, E. (1978). A theorem on utilitarianism. *Review of Economic Studies*, *45*.

Maskin, E. (1979). Decision-making under ignorance with implications for social choice. *Theory and Decision*, *11*.

Meade, J. (1976). *The just economy*. London: Macmillan, and Albany, NY: State University of New York Press.

Moulin, H. (1983). *The strategy of social choice*. Amsterdam: North-Holland.

Nussbaum, M., & Sen, A. (Eds.), (1993). *The quality of life*. Oxford: Oxford University Press.

Paine, T. (1791). *The rights of man: Being an answer to Mr. Burke's attack on the French Revolution*, 1791; republished, *The rights of man*. London: Dent, and New York: Dutton, 1906.

Pattanaik, P. K. (2002). Positional rules of collective decision-making. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *The handbook of social choice and welfare*. Amsterdam: North-Holland.

Pattanaik, P. K., & Salles, M. (Eds.). (1983). *Social choice and welfare*. Amsterdam: North-Holland.

Pattanaik, P. K., & Suzumura, K. (1994). Rights, welfarism, and social choice. *American Economic Review*, *84*.

Peleg, B. (1984). *Game theoretic analysis of voting in committees*. Cambridge: Cambridge University Press.

Phelps, E. S. (Ed.). (1973). *Economic justice*. Harmondsworth: Penguin.

Phelps, E. S. (1977). Recent developments in welfare economics: Justice et équité. In M. D. Intriligator (Eds.), *Frontiers of quantitative economics* (Vol. 3). Amsterdam: North-Holland.

Plott, C. (1978). Rawls' theory of justice: An impossibility result. In H. W. Gottinger & W. Leinfellner (Ed.), *Decision theory and social ethics: Issues in social choice*. Dordrecht: Reidel.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Roberts, K. W. S. (1980a). Interpersonal comparability and social choice theory. *Review of Economic Studies*, *47*.

Roberts, K. W. S. (1980b). Price independent welfare prescriptions. *Journal of Public Economics*, *13*.

Rothschild, E. (2001). *Economic sentiments: Adam Smith, Condorcet and the Enlightenment*. Cambridge, MA: Harvard University Press.

Sen, A. K. (1969). Quasi-transitivity, rational choice and collective decisions. *Review of Economic Studies*, *36*.

Sen, A. K. (1970a). *Collective choice and social welfare*. San Francisco: Holden Day, republished Amsterdam: North-Holland, 1979.

Sen, A. K. (1970b). The impossibility of a Paretian liberal. *Journal of Political Economy*, *78*.

Sen, A. K. (1977a). Social choice theory: A re-examination. *Econometrica*, *45*, reprinted in Sen (1982).

Sen, A. K. (1977b). On weights and measures: informational constraints in social welfare analysis. *Econometrica*, *45*, reprinted in Sen (1982).

Sen, A. K. (1979). Informational analysis of moral principles. In R. Harrison (Ed.), *Rational action*. Cambridge: Cambridge University Press.

Sen, A. K. (1982). *Choice, welfare and measurement*. Oxford: Basil Blackwell, and Cambridge, MA: MIT Press.

Sen, A. K. (1986). Social choice theory. In K. J. Arrow & M. Intriligator (Eds.), *Handbook of mathematical economics*. Amsterdam: North-Holland.

Sen, A. K. (1991). *On indexing primary goods and capabilities*. mimeographed, Harvard University, July 1991.

Sen, A. K. (1993). Internal consistency of choice. *Econometrica*, *61*.

Sen, A. K. (1995). Rationality and social choice. *American Economic Review*, *85*.

Sen, A. K. (1999). The possibility of social choice. *American Economic Review*, *109*.

Sen, A. K. (2009). *The idea of justice*. Cambridge, MA: The Belknap Press of Harvard University Press.

Sen, A., & Williams, B. (Eds.). (1982). *Utilitarianism and beyond*. Cambridge: Cambridge University Press.

Shama, S. R. (1967). *Kautilya's Arthashastra* (8th ed.). Mysore: Mysore Printing and Publishing House.

Smith, A. (1759). *The theory of moral sentiments*. (revised edition, 1790; republished, New York: Penguin, 2009).

Strasnick, S. (1976). Social choice and the derivation of Rawls's Difference Principle. *Journal of Philosophy*, *73*.

Suzumura, K. (1976a). Rational choice and revealed preference. *Review of Economic Studies*, *43*.

Suzumura, K. (1976b). Remarks on the theory of collective choice. *Economica*, *43*.

Suzumura, K. (1982), Equity, efficiency and rights in social choice. *Mathematical Social Sciences*, *3*.

Suzumura, K. (1983). *Rational choice, collective decisions and social welfare*. Cambridge: Cambridge University Press.

Suzumura, K. (1996). Welfare, rights, and social choice procedure: a perspective. *Analyse & Kritik*, *18*.

Vickrey, W. (1945). Measuring marginal utility by reactions to risk. *Econometrica*, *13*.

Wollstonecraft, M. (1790, 1792). *A vindication of the rights of men, in a letter to the right Honourable Edmund Burke; occasions by his reflections on the revolution in France*, 1790, and, *A vindication of the rights of woman: With strictures on political and moral subjects*, 1792; both included in Mary Wollstonecraft, *A vindication of the rights of man and a vindication of the rights of women*, S. Tomaselli (Ed.), Cambridge: Cambridge University Press, 1995.

Young, H. P. (1988). Condorcet's theory of voting. *American Political Science Review*, *82*.

ON THE ROBUSTNESS OF MAJORITY RULE

PARTHA DASGUPTA AND ERIC MASKIN

ABSTRACT

We show that simple majority rule satisfies five standard and attractive axioms—the Pareto property, anonymity, neutrality, independence of irrelevant alternatives, and (generic) decisiveness—over a larger class of preference domains than (essentially) any other voting rule. Hence, in this sense, it is the most robust voting rule. This characterization of majority rule provides an alternative to that of May (1952). (JEL: D71)

1. INTRODUCTION

How should a society select a president? How should a legislature decide which version of a bill to enact?

Reprinted from *Journal of the European Economic Association*, John Wiley and Sons, vol. 6(5), 949–973.

The casual response to these questions is probably to recommend that a vote be taken. But there are many possible voting rules—majority rule, plurality rule, rank-order voting, unanimity rule, runoff voting, and a host of others (a *voting rule*, in general, is any method for choosing a winner from a set of candidates on the basis of voters' reported preferences for those candidates[1])—and so this response, by itself, does not resolve the question. Accordingly, the theory of voting typically attempts to evaluate voting rules systematically by examining which fundamental properties or axioms they satisfy.

One generally accepted axiom is the *Pareto property*, the principle that if all voters prefer candidate x to candidate y, then y should not be chosen over x.[2] A second axiom with strong appeal is *anonymity*, the notion that no voter should have more influence on the outcome of an election than any other[3] (anonymity is sometimes called the "one person–one vote" principle). Just as anonymity demands that all voters be treated alike, a third principle, *neutrality*, requires the same thing for candidates: No candidate should get special treatment.[4]

Three particularly prominent voting rules that satisfy all three axioms—Pareto, anonymity, and neutrality—are (i) *simple majority rule*, according to which candidate x is chosen if, for all other candidates y in the feasible set, more voters prefer x to y than y to x; (ii) *rank-order voting* (also called the Borda count[5]), under which each candidate gets one point for every voter who ranks her first, two points for every voter who ranks her second, and so forth, with candidate *x* being chosen if *x*'s point total is lowest among those in the feasible set; and (iii) *plurality rule* (also called "first past the post"), according to which candidate *x* is chosen if more voters rank *x* first than they do any other feasible candidate.

But rank-order voting and plurality rule fail to satisfy a fourth standard principle, *independence of irrelevant alternatives* (IIA), which has attracted considerable attention since its emphasis by Nash (1950) and Arrow (1951).[6] IIA dictates that if candidate *x* is chosen from the feasible set, and now some other candidate *y* is removed from that set, then *x* is still chosen.[7] To see why rank-order voting and plurality rule violate IIA, consider an electorate consisting of 100 voters. Suppose that there are four feasible candidates *w*, *x*, *y*, *z*, and that the distribution of rankings is as follows:

Number of voters	**47**	**49**	**4**
Ranking (listed vertically from best to worst)	*x*	*y*	*w*
	y	*z*	*x*
	z	*x*	*y*
	w	*w*	*z*

Then, under rank-order voting, *y* will win the election for this profile (a *profile* is a specification of all voters' rankings) with a point total of 155 (49 points from 49 first-place votes, 94 points from 47 second-place votes, and 12 points from 4 third-place votes), compared with point totals of 202 for *x*, 388 for *w*, and 255 for *z*. Candidate *y* will also win under plurality rule: *y* gets 49 first-place votes whereas

x and z get only 47 and 4, respectively. Observe, however, that if z is eliminated from the feasible set, then x will win under rank-order voting with a point total of 153 (47 points from first-place votes and 106 points from second-place votes) compared with y's point total of 155. Moreover, if w is removed from the feasible set (either instead of or in addition to z), then x will also win under plurality rule: It is now top-ranked by 51 voters, whereas y has only 49 first-place votes. Thus, whether the candidates w and z are present or absent from the feasible set determines the outcome under both rank-order voting and plurality rule, contradicting IIA. Furthermore, this occurs even though neither w nor z comes close to winning under either voting rule (i.e., they are "irrelevant alternatives").

Under majority rule (we will henceforth omit the qualification "simple" when this does not cause confusion with other variants of majority rule), by contrast, the choice between x and y turns only on how many voters prefer x to y and how many prefer y to x—not on whether other candidates are also options. Thus, in our 100-voter example, x is the winner (she beats all other candidates in head-to-head comparisons) whether or not w or z is on the ballot. In other words, majority rule satisfies IIA.

But majority rule itself has a well-known flaw, discovered by Borda's archrival the Marquis de Condorcet (1785) and illustrated by the so-called paradox of voting (or Condorcet paradox): it may fail to generate any winner. Specifically, suppose there are three voters 1, 2, 3 and three candidates x, y, z, and suppose the profile of voters' preferences is

1	2	3
x	y	z
y	z	x
z	x	y

(i.e., voter 1 prefers x to y to z, voter 2 prefers y to z to x, and voter 3 prefers z to x to y). Then, as Condorcet noted, a (two-thirds) majority prefers x to y, so that y cannot be chosen; a majority prefers y to z, so that z cannot be chosen; and a majority prefers z to x, so that x cannot be chosen. That is, majority rule fails to select any alternative; it violates *decisiveness*, which requires that a voting rule pick a (unique) winner.

In view of the failure of these three prominent voting methods—rank-order voting, plurality rule, and majority rule—to satisfy all of the five axioms (Pareto, anonymity, neutrality, IIA, and decisiveness), it is natural to inquire whether there is some other voting rule that might succeed where they fail. Unfortunately, the answer is negative: *no* voting rule satisfies all five axioms when there are three or more candidates (see Theorem 1), a result closely related to Arrow's (1951) impossibility theorem.

Still, there is an important sense in which this conclusion is too pessimistic: It presumes that, in order to satisfy an axiom, a voting rule must conform to that axiom regardless of what voters' preferences turn out to be.[8] In practice, however, some preferences may be highly unlikely. One reason for this may be ideology. As Black (1948) notes, in many elections the typical voter's attitudes toward the leading candidates will be governed largely by how far away they are from his own position in left-right ideological space. In the 2000 U.S. presidential election—where the four major candidates from left to right were Ralph Nader, Al Gore, George W. Bush, and Pat Buchanan—a voter favoring Gore might thus have had the ranking

Gore		Gore
Nader	or even	Bush
Bush		Nader
Buchanan		Buchanan

but would most likely not have ranked the candidates as

Gore
Buchanan
Bush
Nader

because Bush is closer to Gore ideologically than Buchanan is. In other words, the graph of a voter's utility for candidates will be *single-peaked* when the candidates are arranged ideologically on the horizontal axis. Single-peakedness is of interest because, as Black shows, majority rule satisfies decisiveness generically[9] when voters' preferences conform to this restriction.

In fact, single-peakedness is by no means the only plausible restriction on preferences that ensures the decisiveness of majority rule. The 2002 French presidential election, where the three main candidates were Lionel Jospin (Socialist), Jacques Chirac (Conservative), and Jean-Marie Le Pen (National Front), offers another example. In that election, voters—regardless of their views on Jospin and Chirac—had strong views on Le Pen: polls suggested that, among the three candidates, he was ranked either first or third by nearly everybody; very few voters placed him second. Whether such polarization is good for France is open to debate, but it is definitely good for majority rule: As we will see in Section 4, such a restriction—in which one candidate is never ranked second—guarantees, like single-peakedness, that majority rule will be generically decisive.

Thus, majority rule *works well*—in the sense of satisfying our five axioms—for some domains of voters' preferences (e.g., a domain of single-peaked preferences), but not for others (e.g., the unrestricted domain). A natural issue to raise, therefore, is how its performance compares with that of other voting rules. As we have already noted, no voting rule can work well for all domains. So the obvious question to ask is: Which voting rules work well for the biggest class of domains?[10]

We show that majority rule is (essentially) the unique answer to this question. Specifically, we establish (see Theorem 2) that, if a given voting rule F works well on a domain of preferences, then majority rule works well on that domain, too. Conversely, if F differs from majority rule,[11] then there exists some other domain on which majority rule works well but F does not.

Thus, majority rule is essentially the unique voting rule that works well on the most domains; it is, in this sense, the most *robust* voting rule.[12] Indeed, this gives us a characterization of majority rule (see Theorem 3) that differs from the classic one derived by May (1952). For the case of two alternatives,[13] May shows that majority rule is the unique voting rule satisfying a weak version of decisiveness, anonymity, neutrality, and a fourth property, *positive responsiveness.*[14] Our Theorem 3 strengthens decisiveness, omits positive responsiveness, and imposes Pareto and IIA to obtain a different characterization.

Theorem 2 is also related to a result obtained in Maskin (1995).[15] Like May, Maskin imposes somewhat different axioms from ours. In particular, instead of decisiveness—which requires that there be a unique winner—he allows for the possibility of multiple winners but insists on *transitivity* (indeed, the same is true of earlier versions of this article; see Dasgupta and Maskin 1998): If x beats y and y beats z, then x should beat z. But more significantly, his proposition requires two strong and somewhat unpalatable assumptions. The first is that the number of voters be odd. This is needed to rule out exact ties: situations where exactly half the population prefers x to y and the other half prefers y to x (oddness is also needed for much of the early work on majority rule; see, e.g., Inada 1969). In fact, our own results also call for avoiding such ties. But rather than simply assuming an odd number of voters, we suppose that the number of voters is large, implying that an exact tie is unlikely even if the number is not odd. Hence, we suppose a large number of voters and ask only for *generic* decisiveness (i.e., decisiveness for "almost all" profiles). Formally, we work with a continuum of voters,[16] but it will become clear that we could alternatively assume a large but finite number by defining generic decisiveness to mean "decisive for a sufficiently high proportion of profiles." In this way we avoid "oddness," an unappealing assumption because it presumably holds only half the time.

Second, Maskin's (1995) proof invokes the restrictive assumption that the voting rule F being compared with majority rule satisfies Pareto, anonymity, IIA, and neutrality on any domain. This is quite restrictive because, although it accommodates certain methods (such as the supermajority rules and the Pareto-extension rule—the rule that chooses all Pareto optimal alternatives), it eliminates such voting rules as the Borda count, plurality voting, and run-off voting. These are the most common alternatives in practice to majority rule, yet they fail to satisfy IIA on the unrestricted domain. We show that this assumption can be dropped altogether.

We proceed as follows. In Section 2, we set up the model. In Section 3, we give formal definitions of our five properties: Pareto, anonymity, neutrality, independence of irrelevant alternatives, and generic decisiveness. We also show (Theorem 1) that no voting rule always satisfies these properties—that is, always works well. In Section 4 we establish three lemmas that characterize when rank-order voting, plurality rule, and majority rule work well. We use the third lemma in Section 5 to establish our main result, Theorem 2. We obtain our alternative to May's (1952) characterization as Theorem 3. Finally, in Section 6 we discuss two extensions.

2. THE MODEL

Our model in most respects falls within a standard social choice framework. Let X be the set of social alternatives (including alternatives that may turn out to be infeasible), which, in a political context, is the set of candidates.

For technical convenience, we take X to be finite with cardinality $m \geq 3$. The possibility of individual indifference often makes technical arguments in the social choice literature a great deal messier (see, e.g., Sen and Pattanaik 1969). We shall simply rule it out by assuming that voters' preferences over X can be represented by *strict orderings*[17] (with only a finite number of alternatives, the assumption that a voter is not exactly indifferent between any two alternatives does not seem very strong). If R is a strict ordering of X, then, for any two alternatives $x, y \in X$ with $x \neq y$, the notation xRy denotes "x is strictly preferred to y in ordering R." For any subset $Y \subseteq X$ and any strict ordering R, let $R|_Y$ be the restriction of R to Y.

Let $\mathfrak{R}_X$ be the set of all logically possible strict orderings of X. We shall typically suppose that voters' preferences are drawn from some subset $\mathfrak{R} \subseteq \mathfrak{R}_X$. For example, for some sequential arrangement $(x_1, x_2, \ldots, x_m)$ of the social alternatives, $\mathfrak{R}$ consists of *single-peaked preferences* (relative to this arrangement) if, for all $R \in \mathfrak{R}$ whenever $x_i R x_{i+1}$ for some i we have $x_j R x_{j+1}$ for all $j > i$ (i.e., if x lies between x_i and x_j in the arrangement, then a voter cannot prefer both x_i and x_j to x).

For the reason mentioned in the Introduction (and elaborated on hereafter), we shall suppose that there is a continuum of voters, indexed by points in the unit interval $[0, 1]$. A *profile* $\boldsymbol{R}$ on $\mathfrak{R}$ is a mapping $\boldsymbol{R}:[0, 1] \rightarrow \mathfrak{R}$, where, for any $i \in [0, 1]$, $\boldsymbol{R}(i)$ is voter i's preference ordering. Hence, profile $\boldsymbol{R}$ is a specification of the preferences of all voters. For any $Y \subseteq X$, $\boldsymbol{R}|_Y$ is the profile $\boldsymbol{R}$ restricted to Y.

We shall use Lebesgue measure μ as our measure of the size of voting blocs. Given alternatives x and y with $x \neq y$ and profile $\boldsymbol{R}$, let $q_{\mathrm{R}}(x, y) = \mu\{i | x\boldsymbol{R}(i)y\}$.[18] Then $q_R(x, y)$ is the fraction of the population preferring x to y in profile $\boldsymbol{R}$.

A *voting rule F* is a mapping that, for each profile[19] $\boldsymbol{R}$ on $\mathfrak{R}_X$ and each subset $Y \subseteq X$, assigns a (possibly empty) subset $F(\boldsymbol{R}, Y) \subseteq X$, where if $\boldsymbol{R}|_Y = \boldsymbol{R}'|_Y$, then $F(\boldsymbol{R}, Y) = F(\boldsymbol{R}', Y)$.[20] As suggested in the Introduction, Y can be interpreted as the *feasible set* of alternatives and $F(\boldsymbol{R}, Y)$ as the *winning candidate(s)*.

For example, suppose that F^M is *simple majority rule*. Then, for all $\boldsymbol{R}$ and Y,

$$F^M(\boldsymbol{R}, Y) = \{x \in Y \mid q_{\mathrm{R}}(x, y) \geq q_{\mathrm{R}}(y, x) \text{ for all } y \in Y - \{x\}\};$$

in other words, x is a winner in Y provided that, for any other alternative $y \in Y$, the proportion of voters preferring x to y is no less than the proportion preferring y to x. Such an alternative x is called a *Condorcet winner*. Note that there may not always be a Condorcet winner—that is, $F^M(\boldsymbol{R}, Y)$ need not be nonempty (as when the profile corresponds to that in the Condorcet paradox).

The supermajority rules provide a second example. For instance, *two-thirds majority* rule $F^{2/3}$ can be defined so that, for all $\boldsymbol{R}$ and Y,

$$F^{2/3}(\boldsymbol{R}, Y) = Y'.$$

Here Y' is a nonempty subset of Y such that, for all $x, y \in Y'$ with $x \neq y$ and all $z \in Y - Y'$, we have $q_R(y, x) < 2/3$ and $q_R(x, z) \geq 2/3$, provided that such a subset exists (if Y' exists, then it is clearly unique). That is, x is a winner if it beats all nonwinners by at least a two-thirds majority and if no other winner beats it by a two-thirds majority or more.

As a third example, consider rank-order voting. Given $R \in \Re_X$ and Y, let $v_R^Y(x)$ be 1 if x is the top-ranked alternative of R in Y, let it be 2 if x is second-ranked in Y, and so on. Then, given profile $\boldsymbol{R}$, it follows that $\int_0^1 v_{\boldsymbol{R}(i)}^Y(x)d\mu(i)$ is the total number of points assigned to x—that is, alternative x's rank-order score or *Borda count*. If F^{RO} is *rank-order voting*, then, for all $\boldsymbol{R}$ and Y,

$$F^{RO}(\boldsymbol{R},Y)=\left\{x\in Y \middle| \int_0^1 v_{\boldsymbol{R}(i)}^Y(x)d\mu(i)\leq\int_0^1 v_{\boldsymbol{R}(i)}^Y(y)d\mu(i),\ \text{for all } y\in Y\right\}.$$

That is, x is a rank-order winner if no other alternative in Y has a lower rank-order score.

Finally, consider plurality rule F^P defined so that, for all $\boldsymbol{R}$ and Y,

$$\begin{aligned}F^P(\boldsymbol{R},Y)=\{x\in Y\,|\,\mu\{i\,|\,x\,\boldsymbol{R}(i)\,y \text{ for all } y\in Y-\{x\}\}\\ \geq\mu\{i\,|\,z\,\boldsymbol{R}(i)\,y \text{ for all } y\in Y-\{z\}\} \text{ for all } z\in Y\}.\end{aligned}$$

That is, x is a plurality winner if it is top-ranked in Y for at least as many voters as any other alternative in Y.

3. THE PROPERTIES

We consider five standard properties that one may wish a voting rule to satisfy.

Pareto Property on $\mathfrak{R}$. For all $\boldsymbol{R}$ on $\mathfrak{R}$ and all $x, y \in X$ with $x \neq y$, if $x\boldsymbol{R}(i)y$ for all i, then, for all Y, $x \in Y$ implies $y \notin F(\boldsymbol{R}, Y)$.

In words, the Pareto property requires that, if all voters prefer x to y, then the voting rule should not choose y if x is feasible. Probably all voting rules used in practice satisfy this property. In particular, majority rule, rank-order voting, and plurality rule (as well as the supermajority rules) satisfy it on the unrestricted domain $\mathfrak{R}_X$.

Anonymity on $\mathfrak{R}$. Suppose that $\pi: [0,1] \rightarrow [0,1]$ is a measure-preserving permutation of $[0, 1]$ (by "measure-preserving" we mean that, for all Borel sets $T \subset [0,1], \mu(T) = \mu(\pi(T))$). If, for all $\boldsymbol{R}$ on $\mathfrak{R}$, $\boldsymbol{R}^{\pi}$ is the profile such that $\boldsymbol{R}^{\pi}(i) = \boldsymbol{R}(\pi(i))$ for all i, then, for all Y, $F(\boldsymbol{R}^{\pi}, Y) = F(\boldsymbol{R}, Y)$.

In words, anonymity means that the winner(s) should not depend on which voter has which preference; only the preferences themselves matter. Thus, if we permute the assignment of voters' preferences by π, the winners should remain the same. (The reason for requiring that π be measure-preserving is purely technical: to ensure that, for all x and y, the fraction of voters preferring x to y is the same for $\boldsymbol{R}^{\pi}$ as it is for $\boldsymbol{R}$.) Anonymity embodies the principle that everyone's vote should count equally.[21] It is obviously satisfied on $\mathfrak{R}_X$ by majority rule, plurality rule, and rank-order voting, as well as by all other voting rules that we have discussed so far.

Neutrality on $\mathfrak{R}$. For any subset $Y \subseteq X$ and profile $\boldsymbol{R}$ on $\mathfrak{R}$, let $\rho: Y \rightarrow Y$ be a permutation of Y and let $\boldsymbol{R}^{\rho,Y}$ be a profile on $\Re$ such that, for all i and all $x,\ y \in Y$ with $x \neq y$, $x\boldsymbol{R}(i)y$ if and only if $\rho(x)\boldsymbol{R}^{\rho,Y}(i)\rho(y)$. Then $\rho(F(\boldsymbol{R},Y)) = F(\boldsymbol{R}^{\rho,Y},Y)$.

In words, neutrality requires that a voting rule treat all alternatives symmetrically: If the alternatives are relabeled via ρ, then the winner(s) are relabeled in the same way. Once again, all the voting rules we have talked about satisfy neutrality, including majority rule, rank-order voting, and plurality rule.

As noted in the Introduction, we will invoke the Nash (1950) version of IIA as follows.

Independence of Irrelevant Alternatives on $\Re$. For all profiles $\boldsymbol{R}$ on $\Re$ and all Y, if $x \in F(\boldsymbol{R},Y)$ and if Y' is a subset of Y such that $x \in Y'$, then $x \in F(\boldsymbol{R},Y')$.

In words, IIA says that if x is a winner for some feasible set Y and we now remove some of the other alternatives from Y, then x will remain a winner. Clearly, majority rule satisfies IIA on the unrestricted domain $\Re_X$: if x beats each other alternative by a majority, then it continues to do so when any of those other alternatives are removed. However, rank-order voting and plurality rule violate IIA on $\Re_X$, as we already showed by example.

Finally, we require that voting rules select a single winner.

Decisiveness. For all $\boldsymbol{R}$ and Y, $F(\boldsymbol{R}, Y)$ is a *singleton*—that is, it consists of a unique element.

Decisiveness formalizes the reasonably uncontroversial goal that an election should result in a clear-cut winner.[22] However, it is somewhat too strong because it rules out ties, even if these occur only rarely. Suppose, say, that $Y = \{x, y\}$

and that exactly half the population prefers x to y and the other half prefers y to x. Then no neutral voting rule will be able to choose between x and y; they are perfectly symmetric in this profile. Nevertheless, this indecisiveness is a knife-edge phenomenon—it requires that the population be split precisely 50–50. Thus, there is good reason for us to disregard it as pathological or irregular. And, because we are working with a continuum of voters, there is a simple formal way to do so.

Specifically, let S be a subset of $\mathbb{R}_+$. A profile $\boldsymbol{R}$ on $\mathfrak{R}$ is *regular* with respect to S (which we call an *exceptional set*) if, for all alternatives x and y with $x \neq y$,

$$q_R(x, y)/q_R(y, x) \notin S.$$

In words, a regular profile is one for which the proportions of voters preferring one alternative to another all fall outside the specified exceptional set. We can now state the version of decisiveness that we will use.

Generic Decisiveness on $\mathfrak{R}$. There exists a finite exceptional set S such that, for all Y and all profiles $\boldsymbol{R}$ on $\mathfrak{R}$ that are regular with respect to S, $F(\boldsymbol{R}, Y)$ is a singleton.

Generic decisiveness requires that a voting rule be decisive for regular profiles, where the preference proportions do not fall into some finite exceptional set. For example, as Lemma 3 implies, majority rule is generically decisive on a domain of single-peaked preferences because there exists a unique winner for all regular profiles if the exceptional set consists of the single point 1 (i.e., $S = \{1\}$). It is this decisiveness requirement that works against such supermajority methods as two-thirds majority rule, which selects a unique winner x only if x beats all other alternatives by at least a two-thirds majority. In fact, in view of the Condorcet paradox, simple majority rule itself is not generically decisive on the domain $\mathfrak{R}_X$. By contrast, rank-order voting and plurality rule are generically decisive on all domains, including $\mathfrak{R}_X$.[23]

We shall say that a voting rule *works well* on a domain $\mathfrak{R}$ if it satisfies the Pareto property, anonymity, neutrality, IIA, and generic decisiveness on that domain. Thus, given our previous discussion, majority rule works well, for example, on a domain of single-peaked preferences. In Section 4 we provide general characterizations of when majority rule, plurality rule, and rank-order voting work well.

Although decisiveness is the only axiom for which we are considering a "generic" version, we could easily accommodate generic relaxations of the other conditions, too. However, this seems pointless, because, to our knowledge, no commonly used voting rule has nongeneric failures except with respect to decisiveness.

We can now establish the impossibility result that motivates our examination of restricted domains $\mathfrak{R}$.

THEOREM 1. *No voting rule works well on* $\mathfrak{R}_X$.

Proof. Suppose, contrary to the claim, that F works well on $\mathfrak{R}_X$. We will use F to construct a social welfare function satisfying the Pareto property, anonymity, and IIA (the Arrow 1951 version), contradicting the Arrow impossibility theorem.

Let S be the exceptional set for F on $\mathfrak{R}$. Because S is finite (by definition of generic decisiveness), we can find an integer $n \geq 2$ such that, if we divide the population into n groups of equal size $[0, 1/n]$, $(1/n, 2/n]$, $(2/n, 3/n]$, . . . , $(n - 1/n, 1]$, then any profile for which all voters within a given group have the same ranking must be regular with respect to S. Given profile $\boldsymbol{R}$ for which all voters within a given group have the same ranking and $X' \subseteq X$, let $\boldsymbol{R}^{X'}$ be the same profile as $\boldsymbol{R}$ except that the elements of X' have been moved to the top of all voters' rankings: for all i and for all $x, y \in X$ with $x \neq y$, $x\boldsymbol{R}^{X'}(i)y$ if and only if

(a) $x\boldsymbol{R}(i)y$ and $x, y \in X'$; or
(b) $x\boldsymbol{R}(i)y$ and $x, y \notin X'$; or
(c) $x \in X'$ and $y \notin X'$.

Construct an n-person social welfare function $f : \mathfrak{R}^n_X \rightarrow \mathfrak{R}_X$ such that, for all n-tuples $(R_1, \ldots, R_n) \in \mathfrak{R}^n_X$ and $x, y \in X$ with $x \neq y$,

$$xf(R_1, \ldots, R_n)y \text{ if and only if } x \in F(\boldsymbol{R}^{\{x,y\}}, X). \qquad (1)$$

Here $\boldsymbol{R}$ corresponds to $(R_1, \ldots, R_n)$; it is the profile such that, for all i and j, $\boldsymbol{R}(i) = R_j$ if and only if $i \in (j/n, j + 1/n]$ (i.e., if voter i belongs to group j). To begin with, f is well defined because, since F satisfies the Pareto Principle and generic decisiveness, either $x \in F(\boldsymbol{R}^{\{x,y\}}, X)$ or $y \in F(\boldsymbol{R}^{\{x,y\}}, X)$. Similarly, f satisfies the Pareto Principle and anonymity.[24] To see that f satisfies Arrow-IIA, consider two n-tuples $(R_1, \ldots, R_n)$ and $(\hat{R}_1, \ldots, \hat{R}_n)$ such that

$$(R_1, \ldots, R_n)|_{\{x,y\}} = (\hat{R}_1, \ldots, \hat{R}_n)|_{\{x,y\}}, \tag{2}$$

and let $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ be the corresponding profiles. From generic decisiveness, Pareto, and IIA, we obtain

$$\begin{aligned} F(\boldsymbol{R}^{\{x,y\}}, X) &= F(\boldsymbol{R}^{\{x,y\}}, \{x,y\} \in \{x,y\}, \\ F(\hat{\boldsymbol{R}}^{\{x,y\}}, X) &= F(\hat{\boldsymbol{R}}^{\{x,y\}}, \{x,y\} \in \{x,y\}. \end{aligned} \tag{3}$$

But from equation (2) and the definition of a voting rule it follows that $F(\boldsymbol{R}^{\{x,y\}}, \{x,y\}) = F(\hat{\boldsymbol{R}}^{\{x,y\}}, \{x,y\})$. Hence, by equations (1) and (3),

$$xf(R_1, \ldots, R_n)y \text{ if and only if } xf(\hat{R}_1, \ldots, \hat{R}_n)y$$

establishing Arrow-IIA.

Finally, we must show that f is transitive. That is, for any n-tuple $(R_1, \ldots, R_n)$ and distinct alternatives x, y, z for which $xf(R_1, \ldots, R_n)y$ and $yf(R_1, \ldots, R_n)z$, we must establish that $xf(R_1, \ldots, R_n)z$. Consider $F(\boldsymbol{R}^{\{x,y,z\}}, X)$, where $\boldsymbol{R}$ is the profile corresponding to $(R_1, \ldots, R_n)$. Because $\boldsymbol{R}^{\{x,y,z\}}$ is regular, generic decisiveness implies that $F(\boldsymbol{R}^{\{x,y,z\}}, X)$ is a singleton, and the Pareto property implies that $F(\boldsymbol{R}^{\{x,y,z\}}, X) \in \{x, y, \mathrm{z}\}$. If $F(\boldsymbol{R}^{\{x,y,z\}}, X) = y$ then, by IIA, $F(\boldsymbol{R}^{\{x,y,z\}}, \{x,y\}) = y$. But from $xf(R_1, \ldots, R_n)y$ and IIA we obtain $F(\boldsymbol{R}^{\{x,y\}}, X) = F(\boldsymbol{R}^{\{x,y\}}, \{x,y\}) = x$; this is a contradiction because, by definition of a voting rule, $F(\boldsymbol{R}^{\{x,y,z\}}, \{x,y\}) = F(\boldsymbol{R}^{\{x,y\}}\ \{x,y\})$. If $F(\boldsymbol{R}^{\{x,y,z\}}, X) = z$ then we can derive a similar contradiction from $yF(R_1, \ldots, R_n)z$. Hence $F(\boldsymbol{R}^{\{x,y,z\}}, X) = x$, and so by definition we have $F(\boldsymbol{R}^{\{x,z\}}, X) = x$, implying that $xF(R_1, \ldots, R_n)\ z$. Thus, transitivity obtains and so f is a social welfare function satisfying Pareto, anonymity, and IIA. The Arrow impossibility theorem now applies to obtain the theorem (anonymity implies that Arrow's nondictatorship requirement is satisfied).

That F satisfies neutrality is a fact not used in the proof, so Theorem 1 remains true if we drop that desideratum from the definition of "working well."

4. CHARACTERIZATION RESULTS

We have seen that rank-order voting and plurality rule violate IIA on $\mathfrak{R}_X$. We now characterize the domains for which they do satisfy this property. For rank-order voting, "quasi-agreement" is the key.

Quasi-Agreement (QA) on $\mathfrak{R}$. Within each triple of distinct alternatives $\{x, y, z\} \subseteq X$, there exists an alternative, say x, that satisfies one of the following three conditions:

(i) for all $R \in \mathfrak{R}$, xR_y and xR_z;
(ii) for all $R \in \mathfrak{R}$, yR_x and zR_x;
(iii) for all $R \in \mathfrak{R}$, either yR_xR_z or zR_xR_y.

In other words, QA holds on domain $\mathfrak{R}$ if, for any triple of alternatives, all voters with preferences in $\mathfrak{R}$ agree on the relative ranking of one of these alternatives: either it is best within the triple, or it is worst, or it is in the middle.

LEMMA 1. *For any domain $\mathfrak{R}$, rank-order voting F^{RO} satisfies IIA on $\mathfrak{R}$ if and only if QA holds on $\mathfrak{R}$.*

REMARK. Of our five principal axioms, rank-order voting violates only IIA on $\mathfrak{R}_X$. Hence, Lemma 1 establishes that rank-order voting works well on $\mathfrak{R}$ if and only if $\mathfrak{R}$ satisfies QA.

See the Appendix for the proof of Lemma 1.

We turn next to plurality rule, for which a condition called *limited favoritism* is needed for IIA.

Limited Favoritism (*LF*) *on* $\mathfrak{R}$. Within each triple of distinct alternatives $\{x, y, z\} \subseteq X$ there exists an alternative, say x, such that for all $R \in \mathfrak{R}$ either yR_x or zR_x.

That is, LF holds on domain $\mathfrak{R}$ if, for any triple of alternatives, there exists one alternative that is never the favorite (i.e., is never top-ranked) for preferences in $\mathfrak{R}$.

LEMMA 2. *For any domain* $\mathfrak{R}$, *plurality rule* F^P *satisfies IIA on* $\mathfrak{R}$ *if and only if LF holds on* $\mathfrak{R}$.

REMARK. Just as QA characterizes when rank-order voting works well, so Lemma 2 shows that LF characterizes when plurality rule works well, because the other four axioms are always satisfied. Indeed, LF also characterizes when a number of other prominent voting rules, such as *runoff voting*,[25] work well.

Proof of Lemma 2. Suppose first that $\mathfrak{R}$ satisfies LF. Consider profile $\boldsymbol{R}$ on $\mathfrak{R}$ and subset Y such that $x \in F^P(\boldsymbol{R}, Y)$ for some $x \in Y$. Then the proportion of voters in $\boldsymbol{R}$ who rank x first among alternatives in Y is at least as big as that for any other alternative. Furthermore, given LF, there can be at most one other alternative that is top-ranked by anyone. That is, x must get a majority of the first-place rankings among alternatives in Y. But clearly x will only increase its majority if some other alternative y is removed from Y. Hence, $F^P(\boldsymbol{R}, Y - \{y\}) = x$, and so F^P satisfies IIA on $\mathfrak{R}$.

Next suppose that domain $\hat{\mathfrak{R}}$ violates LF. Then there exist $\{x, y, z\}$ and $R_x, R_y, R_z \in \mathfrak{R}$ such that, within $\{x, y, z\}$, x is top-ranked for R_x, z is top-ranked for R_z, and $yR_y zR_y x$. Consider profile $\hat{\boldsymbol{R}}$ on $\mathfrak{R}$ such that 40% of voters have ordering R_x, 30% have R_y, and 30% have R_z. Then $x \in F^P(\hat{\boldsymbol{R}}, \{x, y, z\})$, because x is top-ranked by 40% of the population whereas y and z are top-ranked by only 30% each. Suppose now that y is removed from $\{x, y, z\}$. Note that $z \in F^P(\hat{\boldsymbol{R}}, \{x, z\})$, because z is now top-ranked by 60% of voters. Hence, F^P violates IIA on $\hat{\mathfrak{R}}$.

We turn finally to majority rule. We suggested in Section 3 that a single-peaked domain ensures generic decisiveness, and we noted in the Introduction that the same is true when the domain satisfies the property that, for every triple of alternatives, there is one that is never ranked second. But these are only sufficient conditions for generic transitivity; what we want is a condition that is both sufficient and necessary.

To obtain that condition we first note that, for any three alternatives x, y, z, there are six logically possible strict orderings, which can be sorted into two Condorcet "cycles":[26]

$$\begin{array}{ccc|ccc} x & y & z & x & z & y \\ y & z & x & z & y & x \\ z & x & y & y & x & z \end{array}$$

cycle 1	cycle 2

We shall say that a domain $\mathfrak{R}$ satisfies the *no-Condorcet-cycle* property[27] if it contains no Condorcet cycles. That is, for every triple of alternatives, at least one ordering is missing from each of cycles 1 and 2. (More precisely, for each triple $\{x, y, z\}$, there exist no orderings R, R', R'' in $\mathfrak{R}$ that, when restricted to $\{x, y, z\}$, generate cycle 1 or cycle 2.)

LEMMA 3. *Majority rule is generically decisive on domain $\mathfrak{R}$ if and only if $\mathfrak{R}$ satisfies the no-Condorcet-cycle property.*[28]

Proof. If there existed a Condorcet cycle for alternatives $\{x, y, z\}$ in $\mathfrak{R}$, then we could reproduce the Condorcet paradox by taking $Y = \{x, y, z\}$. Hence, the no-Condorcet-cycle property is clearly necessary.

To show that it is also sufficient, we must demonstrate, in effect, that the Condorcet paradox is the only thing that can interfere with majority rule's generic decisiveness. Toward this end, we suppose that F^M is not generically decisive on domain $\mathfrak{R}$. Then, in particular, if $S = \{1\}$ then there must exist Y and a profile $\boldsymbol{R}$ on $\mathfrak{R}$ that is regular with respect to $\{1\}$ but for which $F^M(\boldsymbol{R}, Y)$ is either empty or contains multiple alternatives. If there exist $x, y \in F^M(\boldsymbol{R}, Y)$ with $x \neq y$, then $q_R(x, y) = q_R(y, x) = 1/2$ and so

$$q_R(x, y)/q_R(y, x) = 1,$$

contradicting $\boldsymbol{R}$'s regularity with respect to $\{1\}$. Hence $F^M(\boldsymbol{R}, Y)$ must be empty. Choose $x_1 \in Y$. Then, because $x_1 \notin F^M(\boldsymbol{R}, Y)$, there exists an $x_2 \in Y$ such that

$$q_R(x_2, x_1) > \frac{1}{2}.$$

Similarly, because $x_2 \notin F^M(\boldsymbol{R}, Y)$, there exists an $x_3 \in Y$ such that

$$q_R(x_3, x_2) > \frac{1}{2}.$$

Continuing in this way, we must eventually (because there are only finitely many alternatives in X) reach $x_t \in Y$ such that

$$q_R(x_t, x_{t-1}) > \frac{1}{2} \tag{4}$$

but with some $t < t$ for which

$$q_R(x_\tau, x_t) > \frac{1}{2}. \tag{5}$$

If t is the smallest index for which (5) holds, then

$$q_R(x_{t-1}, x_\tau) > \frac{1}{2}. \tag{6}$$

Combining (4) and (6), we conclude that there must be a positive fraction of voters in $\boldsymbol{R}$ who prefer x_t to x_{t-1} and x_{t-1} to x_τ; that is,

$$\begin{array}{l} x_t \\ x_{t-1} \in \mathfrak{R}.\text{[29]} \\ x_\tau \end{array} \tag{7}$$

Similarly, (5) and (6) yield

$$\begin{array}{l} x_{t-1} \\ x_\tau \in \mathfrak{R}, \\ x_t \end{array}$$

and from (4) and (5) we obtain

$$\begin{array}{l} x_\tau \\ x_t \in \mathfrak{R}. \\ x_{t-1} \end{array}$$

Hence, $\mathfrak{R}$ violates the no-Condorcet-cycle property, as was to be shown. □

It is easy to see that a domain of single-peaked preferences satisfies the no-Condorcet-cycle property. Hence, Lemma 3 implies that majority rule is generically decisive on such a domain. The same is true of the domain we considered in the Introduction in connection with the 2002 French presidential election.

The results of this section give us an indication of the stringency of the requirement of "working well" across our three voting rules. Lemma 1 establishes that, for any triple of alternatives, four of the six possible strict orderings must be absent from a domain $\mathfrak{R}$ in order for rank-order voting to work well on $\mathfrak{R}$. By contrast, Lemmas 2 and 3 show that only two orderings must be absent if we instead consider plurality rule or majority rule (although LF is strictly a more demanding condition than the no-Condorcet-cycle property[30]).

5. THE ROBUSTNESS OF MAJORITY RULE

We can now state our main finding as follows.

THEOREM 2. *Suppose that voting rule F works well on domain $\mathfrak{R}$. Then majority rule F^M works well on $\mathfrak{R}$ too. Conversely, suppose that F^M works well on domain $\mathfrak{R}^M$. Then, if there exists a profile $\boldsymbol{R}^\circ$ on $\mathfrak{R}^M$, regular with respect to F's exceptional set, such that*

$$F(\boldsymbol{R}^\circ, Y) \neq F^M(\boldsymbol{R}^\circ, Y) \text{ for some } Y, \qquad (8)$$

then there exists a domain $\mathfrak{R}'$ on which F^M works well but F does not.

Remark 1. Without the requirement that the profile $\boldsymbol{R}^\circ$ for which F and F^M differ belong to a domain on which majority rule works well, the second assertion of Theorem 2 would be false. In particular, consider a voting rule that coincides with majority rule except for profiles that violate the no-Condorcet-cycle property. It is easy to see that such a rule works well on any domain for which majority rule does because it coincides with majority rule on such a domain.

Remark 2. Theorem 2 allows for the possibility that $\mathfrak{R}' = \mathfrak{R}^M$, and indeed this equality holds in the example we consider after the proof. However, more generally, F may work well on $\mathfrak{R}^M$ even though equation (8) holds, in which case $\mathfrak{R}'$ and $\mathfrak{R}^M$ must differ.

Proof of Theorem 2. Suppose first that F works well on $\mathfrak{R}$. If, contrary to the theorem, F^M does not work well on $\mathfrak{R}$, then by Lemma 3 there exists a Condorcet cycle in $\mathfrak{R}$:

$$\begin{matrix} x & y & z \\ y & z & x \\ z & x & y \end{matrix} \in \mathfrak{R} \tag{9}$$

for some $x, y, z \in X$. Let S be the exceptional set for F on $\mathfrak{R}$. Because S is finite (by definition of generic transitivity), we can find an integer n such that, if we divide the population into n equal groups, then any profile for which all voters in each particular group have the same ordering in $\mathfrak{R}$ must be regular with respect to S.

Let $[0, 1/n]$ be group 1, let $(1/n, 2/n]$ be group 2, . . . , and let $(n - 1/n, 1]$ be group n. Consider a profile $R1$ on $\Re$ such that all voters in group 1 prefer x to z and all voters in the other groups prefer z to x. That is, the profile is

$$\begin{array}{cccc} \underline{1} & \underline{2} & \cdots & \underline{n} \\ x & z & & z. \\ z & x & & x \end{array} \tag{10}$$

Because F is generically decisive on $\Re$ and because $\boldsymbol{R}_1$ is regular, there are two cases: $F(\boldsymbol{R}_1, \{x, z\}) = z$ or x.

Case (i)*:* $F(\boldsymbol{R}_1, \{x, z\}) = z$. Consider a profile $\boldsymbol{R}_1^*$ on $\Re$ in which all voters in group 1 prefer x to y to z, all voters in group 2 prefer y to z to x, and all voters in the remaining groups prefer z to x to y. That is,[31]

$$\boldsymbol{R}_1^* = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n}. \\ x & y & z & & z \\ y & z & x & & x \\ z & x & y & & y \end{array} \tag{11}$$

By (9), such a profile exists on $\Re$. Notice that, in profile $\boldsymbol{R}_1^*$, voters in group 1 prefer x to z and that all other voters prefer z to x. Hence, the case (i) hypothesis implies that

$$F(\boldsymbol{R}_1^*, \{x, z\}) = z. \tag{12}$$

From equation (12) and IIA, $F(\boldsymbol{R}_1^*, \{x, y, z\}) \neq x$. If $F(\boldsymbol{R}_1^*, \{x, y, z\}) = y$, then neutrality implies that

$$F(\hat{\boldsymbol{R}}_1^*, \{x, y, z\}) = x \tag{13}$$

where

$$\hat{R}_1^* = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n} \\ z & x & y & & y \\ x & y & z & & z \\ y & z & x & & x \end{array}.$$

and $\hat{R}_1^*$ is on $\Re$. Then IIA yields $F(\hat{R}_1^*, \{x, z\}) = x$, which by anonymity, contradicts the case (i) hypothesis. Hence, $F(R_1^*, \{x, y, z\}) = z$ and so, by IIA,

$$F(R_1^*, \{y, z\}) = z. \tag{14}$$

Applying neutrality, we obtain from equation (14) that

$$F(R_2, \{y, z\}) = y,$$

where

$$R_2 = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n} \\ z & z & y & & y \\ x & x & z & & z \\ y & y & x & & x \end{array} \tag{15}$$

and R_2 is on $\Re$. Applying neutrality once again then gives

$$F(\hat{R}_2, \{x, z\}) = z, \tag{16}$$

where

$$\hat{R}_2 = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n} \\ x & x & z & & z \\ y & y & x & & x \\ z & z & y & & y \end{array} \tag{17}$$

and $\hat{R}_2$ is on $\mathfrak{R}$. Formulas (16) and (17) establish that if z is chosen over x when just one of n groups prefers x to z (case (i) hypothesis), then z is again chosen over x when two of n groups prefer x to z as in (16).

Now, choose R_2^* on $\mathfrak{R}$ so that

$$R_2^* = \begin{array}{cccccc} \underline{1} & \underline{2} & \underline{3} & \underline{4} & \cdots & \underline{n}. \\ x & y & y & z & & z \\ y & z & z & x & & x \\ z & x & x & y & & y \end{array}$$

Arguing as we did for R_1^*, we can show that $F(R_2^*, \{y, z\}) = z$ and then apply neutrality twice to conclude that z is chosen over x if three groups out of n prefer x to z. Continuing iteratively, we conclude that z is chosen over x even if $n - 1$ groups out of n prefer x to z—which, in view of neutrality, violates the case (i) hypothesis. Hence, this case is impossible.

Case (ii): $F(R_1, \{x, z\}) = x$. From the case (i) argument, case (ii) leads to the same contradiction as before. We conclude that F^M must work well on $\mathfrak{R}$ after all, as claimed.

For the converse, suppose that there exist (a) domain $\mathfrak{R}^M$ on which F^M works well and (b) Y and $x, y \in Y$ and regular profile R° on $\mathfrak{R}^M$ such that

$$y = F(R^\circ, Y) \neq F^M(R^\circ, Y) = x. \tag{18}$$

If F does not work well on $\mathfrak{R}^M$, then we can take $\mathfrak{R}' = \mathfrak{R}^M$ to complete the proof. Hence, assume that F works well on $\mathfrak{R}^M$ with exceptional set S. From IIA and equation (17) (and because $\boldsymbol{R}^\circ$ is regular), there exists an $\alpha \in (0,1)$ with $\alpha/(1-\alpha) \notin S, (1-\alpha)/\alpha \notin S$,

$$1-\alpha > \alpha, \tag{19}$$

and $q_{R^\circ}(x,y) = 1-\alpha$ such that $F^M(\boldsymbol{R}^\circ, \{x,y\}) = x$ and

$$F(\boldsymbol{R}^\circ, \{x,y\}) = y. \tag{20}$$

Consider $z \notin \{x, y\}$ and profile $\boldsymbol{R}^{\circ\circ}$ such that

$$\boldsymbol{R}^{\circ\circ} = \begin{array}{ccc} \underline{[0,\alpha)} & \underline{[\alpha,1-\alpha)} & \underline{[1-\alpha,1]} \\ z & z & x \\ y & x & z \\ x & y & y \end{array}. \tag{21}$$

Observe that in equation (21) we have left out the alternatives other than x, y, and z. To make matters simple, assume that the orderings of $\boldsymbol{R}^{\circ\circ}$ are all the same for those other alternatives. Suppose, furthermore, that in these orderings x, y, and z are each preferred to any alternative not in $\{x, y, z\}$. Then, because $\alpha/(1-\alpha) \notin / S$ and $(1-\alpha)/\alpha \notin S$, it follows that $\boldsymbol{R}^{\circ\circ}$ is regular.

Let $\hat{\Re}'$ consist of the orderings in $\boldsymbol{R}^{\circ\circ}$ together with ordering $\begin{array}{c} x \\ y \\ z \end{array}$ (where $\{x, y, z\}$ are ranked at the top and the other alternatives are ranked as in the other three orderings). By Lemma 3, F^M works well on $\hat{\Re}'$ so we can assume that F does, too (otherwise, we are done). Given generic decisiveness and that $\boldsymbol{R}^{\circ\circ}$ is regular, $F(\boldsymbol{R}^{\circ\circ}, \{x,y,z\})$ is a singleton. We cannot have $F(\boldsymbol{R}^{\circ\circ}, \{x,y,z\}) = y$, because z Pareto dominates y.

If $F(\boldsymbol{R}^{\circ\circ}, \{x, y, z\}) = x$, then $F(\boldsymbol{R}^{\circ\circ}, \{x, y\}) = x$ by IIA. But anonymity and (21) yield,

$$F(\boldsymbol{R}^{\circ\circ}, \{x, y\}) = F(\boldsymbol{R}^{\circ}, \{x, y\}) = y, \tag{22}$$

a contradiction. Thus, we must have $F(\boldsymbol{R}^{\circ\circ}, \{x, y, z\}) = z$, implying from IIA that $F(\boldsymbol{R}^{\circ\circ}, \{x, z\}) = z$. Then neutrality in turn implies that

$$F(\hat{\boldsymbol{R}}^{\circ\circ}, \{x, z\}) = x, \tag{23}$$

where $\hat{\boldsymbol{R}}^{\circ\circ}$ is a profile on $\hat{\Re}'$ such that

$$\hat{\boldsymbol{R}}^{\circ\circ} = \begin{array}{ccc} \underline{[0,\alpha)} & \underline{[\alpha, 1-\alpha)} & \underline{[1-\alpha, 1]} \\ x & x & z \\ y & z & x \\ z & y & y \end{array}. \tag{24}$$

Next, take $\Re'$ to consist of the orderings in equation (24) together with $\begin{matrix} x \\ y \\ z \end{matrix}$. Again, F^M works well on $\Re'$ and so we can assume that F does, too. By equation (23) by we can deduce, using neutrality, that

$$F(\boldsymbol{R}^{\circ\circ\circ}, \{x, y\}) = x$$

for profile $\boldsymbol{R}^{\circ\circ\circ}$ on $\Re'$ such that

$$\boldsymbol{R}^{\circ\circ\circ} = \begin{array}{ccc} \underline{[0,\alpha)} & \underline{[\alpha, 1-\alpha)} & \underline{[1-\alpha, 1]} \\ x & x & y \\ z & y & x \\ y & z & z \end{array},$$

which, from anonymity, contradicts equation (22). Hence, F does not work well on $\Re'$ after all.

As a simple illustration of Theorem 2, let us see how it applies to rank-order voting and plurality rule. For $X = \{w, x, y, z\}$, Lemmas 1 and 2 imply that F^{RO} and F^{P} work well on the domain

$$\mathfrak{R} = \left\{ \begin{matrix} x & z \\ y, & y \\ z & x \\ w & w \end{matrix} \right\}$$

because $\mathfrak{R}$ satisfies both QA and LF. Moreover, as Theorem 2 guarantees, F^{M} also works well on this domain, because it obviously does not contain a Condorcet cycle.

Conversely, on the domain

$$\mathfrak{R}' = \left\{ \begin{matrix} x & y & w \\ y, & z, & x \\ z & x & y \\ w & w & z \end{matrix} \right\} \tag{25}$$

we have

$$\begin{aligned} x = F^{M}(\boldsymbol{R}^{\circ}, \{w, x, y, z\}) \neq F^{RO}(\boldsymbol{R}^{\circ}, \{w, x, y, z\}) \\ = F^{P}(\boldsymbol{R}^{\circ}, \{w, x, y, z\}) = y \end{aligned}$$

for the profile $\boldsymbol{R}^{\circ}$, as in the Introduction, in which the distribution of rankings is as follows:

Proportion of voters	**0.47**	**0.49**	**0.04**
Ranking	x	y	w
	y	z	x
	z	x	y
	w	w	z

By Lemma 3, F^M works well on $\mathfrak{R}'$ defined by equation (25). Moreover, by Lemmas 1 and 2, F^{RO} and F^P do not work well $\mathfrak{R}'$. Hence, we have an example of why Theorem 2 applies to plurality rule and rank-order voting.

In the Introduction we mentioned May's (1952) axiomatization of majority rule. In view of Theorem 2, we can provide an alternative characterization. Specifically, call two voting rules F and F' *generically identical* on domain $\mathfrak{R}$ if there exists a finite set $S \subset \mathbb{R}_+$ such that $F(\boldsymbol{R}, Y) = F'(\boldsymbol{R}, Y)$ for all Y and all $\boldsymbol{R}$ on $\mathfrak{R}$ for which $q_R(x, y)/q_R(y, x) \notin S$ for all $x, y \in Y$. Call F *maximally robust* if there exists no other voting rule that (a) works well on every domain on which F works well and (b) works well on some domain on which F does not work well. Theorem 2 implies that majority rule can be characterized as essentially the unique voting rule that satisfies Pareto, anonymity, neutrality, IIA, and generic decisiveness on the most domains.

Theorem 3. *Majority rule is essentially the unique maximally robust voting rule* (*Any other maximally robust voting rule F is generically identical to majority rule on any domain on which F or majority rule works well.*)[32]

6. FURTHER WORK

We noted in footnote 7 that IIA is related to the demand that a voting rule should be immune to strategic voting. In a follow-up paper (Dasgupta and Maskin 2007a), we explicitly replace IIA by this requirement of strategic immunity.

In another line of work, we drop the neutrality axiom. The symmetry inherent in neutrality is often a reasonable and desirable property—we would presumably want to treat all candidates in a presidential election the same. However, there are also circumstances in which it is natural to favor particular alternatives. The rules for amending the U.S. Constitution are a case in point. They have been deliberately devised so that, at any time, the current version of the Constitution—the status quo—is difficult to change.

In Dasgupta and Maskin (2007b) we show that, if neutrality is dropped (and the requirement that ties be broken "consistently" is also imposed), then unanimity rule with an order of precedence[33] supplants majority rule as the most robust voting rule (clearly, this version of unanimity rule is highly nonneutral). It is not surprising that, with fewer axioms to satisfy, there should be voting rules that satisfy them all on a wider class of domains than majority rule does. Nevertheless, it is notable that, once again, the maximally robust rule is simple and familiar.

APPENDIX

LEMMA 1. *For any domain* $\mathfrak{R}$, F^{RO} *satisfies IIA on* $\mathfrak{R}$ *if and only if QA holds on* $\mathfrak{R}$.

Proof. Because F^{RO} is generically decisive on $\mathfrak{R}_X$, we can restrict attention to profiles $\boldsymbol{R}$ and subsets Y for which $F^{RO}(\boldsymbol{R}, Y)$ is a singleton. Assume first that QA holds on $\mathfrak{R}$.

We must show that F^{RO} satisfies IIA on $\mathfrak{R}$. Consider profile $\boldsymbol{R}$ on $\mathfrak{R}$ and subset Y such that

$$F^{RO}(\boldsymbol{R}, Y) = x \tag{A.1}$$

for some $x \in Y$. We must show that, for all $y \in Y - \{x\}$,

$$F^{RO}(\boldsymbol{R}, Y - \{y\}) = x. \tag{A.2}$$

Suppose, to the contrary, that

$$F^{RO}(\boldsymbol{R}, Y - \{y\}) = z \text{ for } z \in Y - \{x\}. \tag{A.3}$$

By definition of F^{RO}, equations (A.1) and (A.3) together imply that the deletion of y causes z to rise relative to x in some voters' rankings in $\boldsymbol{R}$—in other words, that those voters have the ranking $\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}$. Hence,

$$\begin{matrix} x \\ y \\ z \end{matrix} \in \mathfrak{R}|_{\{x,y,z\}}. \tag{A.4}$$

But (A.3) also implies that there exists an i such that

$$z\boldsymbol{R}(i)x. \tag{A.5}$$

Therefore, equations (A.4), (A.5), and QA imply that

$$\left\{ \begin{matrix} x \\ y, \\ z \end{matrix} \quad \begin{matrix} z \\ y \\ x \end{matrix} \right\} = \mathfrak{R}|_{\{x,y,z\}}. \tag{A.6}$$

Now, the definition of F^{RO} together with equations (A.1), (A.3), and (A.6) imply that $q_R(x, z) > q_R(z, x)$ (the deletion of y must hurt x's score relative to z more than it helps).

From this inequality it follows that there exist $w \in Y$ and $R \in \Re$ with $zR_y R_x$ and $zR_w R_x$ as well as $R' \in \Re$ with $xR'_y R'_z$ and either (a) wR'_x or (b) zR'_w; otherwise, equation (A.2) will hold. If (a) holds then

$$\left\{\begin{matrix} z & w \\ w, & x \\ x & z \end{matrix}\right\} \subseteq \Re|_{\{x,w,z\}}, \qquad \text{(A.7)}$$

and if (b) holds then

$$\left\{\begin{matrix} z & x \\ w, & z \\ x & w \end{matrix}\right\} \subseteq \Re|_{\{x,y,z\}}. \qquad \text{(A.8)}$$

But equations (A.7) and (A.8) both violate QA, so equation (A.2) must hold after all.

Next, suppose that QA does not hold on $\Re$. Then there exist alternatives x, y, z such that

$$\left\{\begin{matrix} x & y \\ y, & z \\ z & x \end{matrix}\right\} \subseteq \Re|_{\{x,y,z\}}. \qquad \text{(A.9)}$$

Consider the profile $\boldsymbol{R}^*$ in which proportion .6 of the population has ranking $\begin{matrix} x \\ y \\ z \end{matrix}$ and proportion .4 has $\begin{matrix} y \\ z \\ x \end{matrix}$. Then

$$F^{RO}(\boldsymbol{R}^*, \{x, y, z\}) = y.$$

But

$$F^{RO}(\boldsymbol{R}^*, \{x, y\}) = x,$$

contradicting IIA. Therefore, F^{RO} does not work well on $\mathfrak{R}$, as was to be shown. □

NOTES

Acknowledgments: This work was supported by grants from the Beijer International Institute of Ecological Economics, Stockholm, and the National Science Foundation. We thank Salvador Barberà and John Weymark for helpful comments.
E-mail addresses: Dasgupta: partha.dasgupta@econ.cam.ac.uk; Maskin: maskin@ias.edu

1. In many electoral systems, a voter reports only his or her favorite candidate, rather than expressing a ranking of all candidates. If there are just two candidates (as in referenda, where the "candidates" are typically "yes" and "no"), then both sorts of reports amount to the same thing. But with three or more candidates, knowing just the voters' favorites is not enough to conduct some of the most prominent voting methods, such as majority rule and rank-order voting.
2. Although the Pareto property is quite uncontroversial in the context of political elections, it is not always so readily accepted—at least by noneconomists—in other social choice settings. Suppose, for example, that the "candidates" are two different national health care plans. Then, some would argue that factors such as fairness, scope of choice, and degree of centralization should to some degree supplant citizens' preferences in determining the choice between the two plans.
3. It is because the Electoral College violates anonymity—voters from large states do not have the same power as those from small states—that many have called for its abolition in U.S. presidential elections. However, like the Pareto property, anonymity is not always so widely endorsed in non-election settings. In our health care scenario (see footnote 2), for example,

it might be considered proper to give more weight to citizens with low incomes.

4. Neutrality is hard to quarrel with in the setting of political elections. But if instead the "candidates" are, say, various amendments to a nation's constitution, then one might want to give special treatment to the status quo—namely, to *no* change—and so ensure that constitutional change occurs only with overwhelming support.
5. It is called this after the eighteenth-century French engineer, Jean-Charles Borda, who first formalized the rank-order voting rule.
6. The Nash and Arrow versions of IIA differ somewhat. Here we follow the Nash formulation.
7. Independence of irrelevant alternatives—although more controversial than the other three principles—has at least two strong arguments in its favor. First, as the name implies, it ensures that the outcome of an election will be unaffected by whether or not candidates with no chance of winning are on the ballot. Second, IIA is closely connected with the property that voters should have no incentive to vote *strategically*—that is, at variance with their true preferences (see Theorem 4.73 in Dasgupta, Hammond, and Maskin 1979). Still, it has generated considerably more controversy than the other properties, particularly among proponents of rank-order voting (the Borda count), which famously violates IIA (see the text to follow).
8. This is the *unrestricted domain* requirement.
9. We clarify what we mean by "generic" decisiveness in Section 3.
10. It is easy to exhibit voting rules that satisfy four of the five properties on all domains of preferences. For instance, *supermajority rules* such as two-thirds majority rule, which chooses alternative x over alternative y if x garners at least a two-third's majority over y (see Section 2 for a more precise definition), satisfy Pareto, anonymity, neutrality, and IIA on any domain. Similarly, rank-order voting satisfies Pareto, anonymity, neutrality, and (generic) decisiveness on any domain.

11. More accurately, the hypothesis is that F differs from majority rule for a "regular" preference profile (see Section 3) belonging to a domain on which majority rule works well.
12. More precisely, any other maximally robust voting rule can differ from majority rule only for "irregular" profiles on any domain on which it works well (see Theorem 3).
13. May considers only the case of two alternatives, but one can impose IIA (the Arrow 1951 version) and thereby readily obtain an extension to three or more alternatives (see Campbell 1982, 1988; Maskin 1995).
14. A voting rule is *positively responsive* if, wherever alternative x is chosen (perhaps not uniquely) for a given profile of voters' preferences and those preferences are then changed only so that x moves up in some voter's ranking, then x becomes uniquely chosen.
15. See Campbell and Kelly (2000) for a generalization of Maskin's result.
16. To our knowledge, this is the first voting theory report to use a continuum in order to formalize the concept of an axiom being satisfied for almost all profiles.
17. Formally, a *strict ordering* (sometimes called a "linear ordering") is a binary relation that is reflexive, complete, transitive, and antisymmetric (antisymmetry means that if xR_y and $x \neq y$, then it is not the case that yR_x).
18. Because Lebesgue measure is not defined for all subsets of $[0, 1]$, we will restrict attention to profiles $\boldsymbol{R}$ such that, for all x and y, $\{i | x\, \boldsymbol{R}(i) y\}$ is a Borel set. Call these *Borel profiles.*
19. Strictly speaking, we must limit attention to Borel profiles (see footnote 18) but henceforth we will not explicitly state this qualification.
20. The requirement that $F(\boldsymbol{R}, Y) = F(\boldsymbol{R}', Y)$ if $\boldsymbol{R}|_Y = \boldsymbol{R}'|_Y$ may seem to resemble IIA, but it is actually much weaker. It merely says that, given the set of feasible candidates Y, the winner(s) should be determined only by voters' preferences over this set and not

by their preferences for infeasible candidates. Indeed, all the voting rules we have discussed—including rank-order voting and plurality rule—satisfy this requirement.

21. Indeed, it is sometimes called "voter equality" (see Dahl 1989).
22. In some public decision-making settings, the possibility of multiple winning alternatives would not be especially problematic (one could simply randomize among them to make a decision), but in a political election such multiplicity would clearly be unsatisfactory.
23. If $m = 3$, then rank-order voting is generically decisive on $\Re_X$ with exceptional set $S = \{1/2, 1, 2\}$, whereas plurality rule is generically decisive on $\Re_X$ with exceptional set $S = \{1\}$.
24. We have previously defined the Pareto property and anonymity for voting rules. Here we mean their natural counterparts for social welfare functions. Thus, Pareto requires that if everyone prefers *x* to *y* then the social ranking will prefer *x* to *y*, and anonymity dictates that if we permute the rankings in the *n*-tuple, then the social ranking remains the same.
25. *Runoff voting*, which is used for presidential elections in many countries, chooses the plurality winner if that candidate is top-ranked by a majority. Otherwise, it chooses the majority winner in a contest between the two candidates that are top-ranked by the most voters (i.e., there is a runoff between those two).
26. We call these *Condorcet* cycles because they constitute preferences that give rise to the Condorcet paradox.
27. Sen (1966) introduces an equivalent condition and calls it *value restriction.*
28. For the case of a finite and odd number of voters, Inada (1969) establishes that a condition equivalent to the no-Condorcet-cycle property is necessary and sufficient for majority rule to be transitive.
29. To be precise, formula (7) says that there exists an ordering in $R \in \Re$ such that $x_t R_{x_{t-1}} R_{x_\tau}$.

30. To see this, notice that LF rules out Condorcet cycles and that the domain $\left\{\begin{matrix} x & y & y & z \\ y & z & x & y \\ z & x & z & x \end{matrix}\right\}$ violates LF but contains no Condorcet cycle.
31. This is not a complete specification of $\boldsymbol{R}_1^*$ because we are not indicating how voters rank alternatives other than *x*, *y*, and *z*. However, from IIA it follows that these other alternatives do not matter for the argument.
32. Theorem 3 requires the imposition of all five properties: Pareto, anonymity, neutrality, IIA, and generic decisiveness. Without Pareto, *minority rule* (where *x* is chosen if fewer voters prefer *x* to *y* than *y* to *x* for all *y*) is as robust as majority rule. Without anonymity, a *dictatorship* (where choices are made according to the preference ranking of a particular voter, the dictator) is maximally robust because it satisfies the remaining conditions on $\Re_X$, the unrestricted domain. Without neutrality, *unanimity rule with an order ofprecedence* (the rule according to which *x* is chosen over *y* if it precedes *y* in the order of precedence, unless everybody prefers *y* to *x)* becomes maximally robust. Without IIA, rank-order voting and plurality rule both are maximally robust because they satisfy the remaining conditions on $\Re_X$. Finally, without generic decisiveness, the supermajority rules are equally as robust as majority rule.
33. For discussion of this voting rule in a political setting, see Buchanan and Tullock (1962).

REFERENCES

Arrow, Kenneth J. (1951). *Social Choice and Individual Values.* John Wiley.

Black, David (1948). "On the Rationale of Group Decision-making." *Journal of Political Economy*, 56(1), 23–34.

Buchanan, James, and Gordon Tullock (1962). *The Calculus of Consent.* University of Michigan Press.

Campbell, Donald (1982). "On the Derivation of Majority Rule." *Theory and Decision*, 14, 133–140.

Campbell, Donald (1988). "A Characterization of Simple Majority Rule for Restricted Domain." *Economics Letters*, 28, 307–310.

Campbell, Donald, and Jerry Kelly (2000). "A Simple Characterization of Majority Rule." *Economic Theory*, 15, 689–700.

Condorcet, Marie Jean Antoine Nicolas Caritat (1785). *Essai sur l'application de l'analyse* à *la pluralité des voix.* Imprimerie Royale.

Dahl, Robert (1989). *Democracy and Its Critics.* Yale University Press.

Dasgupta, Partha, Peter Hammond, and Eric Maskin (1979). "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility." *Review ofEconomic Studies*, 46(2), 185–216.

Dasgupta, Partha, and Eric Maskin (1998). "On the Robustness of Majority Rule." Working paper, University of Cambridge and Harvard University.

Dasgupta, Partha, and Eric Maskin (2007a). "Majority Rule and Strategic Voting," Working paper, Institute for Advanced Study, School of Social Science.

Dasgupta, Partha, and Eric Maskin (2007b). "On the Robustness of Unanimity Rule." Working paper, Institute for Advanced Study, School of Social Science.

Inada, Ken-Ichi (1969). "On the Simple Majority Rule Decision." *Econometrica*, 37(3), 490–506.

Maskin, Eric (1995). "Majority Rule, Social Welfare Functions, and Game Form." In Kaushik Basu, Prasanta K. Pattanaik, and Kōtarō Suzumura, eds., *Choice, Welfare and Development.* Clarendon Press.

May, Kenneth (1952). "Set of Necessary and Sufficient Conditions for Simple Majority Decisions." *Econometrica*, 20(4), 680–684.

Nash, John F. (1950). "The Bargaining Problem." *Econometrica*, 18(2), 155–162.

Sen, Amartya (1966). "A Possibility Theorem on Majority Decisions." *Econometrica*, 34(2), 491–499.

Sen, Amartya, and Prasanta Pattanaik (1969). "Necessary and Sufficient Conditions for Rational Choice under Majority Decision." *Journal of Economic Theory*, 1(2), 178–202.

THE ORIGINS OF THE IMPOSSIBILITY THEOREM

KENNETH J. ARROW

As requested, this is a personal account of the steps by which I came to state and prove the so-called impossibility theorem for social choices. It will be worthwhile to restate the theorem in abbreviated form in order to understand better its genesis.

We consider a society which must make a choice binding on all its members. On any given occasion, there is a set of alternatives, S, among which the social choice must be made. It will be assumed that both the society and any set of alternatives considered are finite. The ith individual in the society has a preference ordering, R_i, over all conceivable alternatives. Each ordering is assumed to be transitive and complete (that is, for any two alternatives, x and y, at least one of the

Reprinted from Arrow, K. J. (1991). The origins of the impossibility theorem. In J. K. Lenstra, A. H. G. Rinnooy Kan, A. Schrijver (Eds.), *History of mathematical programming: A collection of personal reminiscences*. Amsterdam: North-Holland.

two statements, $x R_{iy}$ or $y R_{ix}$, holds). We seek a social preference ordering, R, over all alternatives from which the social choices for any given set of alternatives, S, can be derived as the maximal elements of S under the ordering R.

The social ordering must satisfy two properties: it must reflect in some sense the preference orderings of individuals, and in making a social choice from any given set of alternatives, it should use information about the preference orderings of individuals among those alternatives only. Further, it should be defined for any conceivable set of individual preference orderings, i.e., it is a functional, called the social welfare functional, from vectors of individual preference orderings to preference orderings. The second property above, termed Irrelevance of Independent Alternatives, has the particular implication that the choice from any two-alternative set depends only on the preferences of individuals as between those alternatives. The first property can be stated in differing ways; a particularly weak one is called the Pareto Principle and states that if everyone strictly prefers x to y, then x is preferred to y in the social ordering. The impossibility theorem then states that there is no social welfare functional with the required properties.

The example that most clearly illustrates the intended scope of the theorem in real-life situations is that of an election. Any form of election uses as data the preferences of the individual voters among the candidates (indeed, usually much less data are available, for example, only the first choice of each voter) and nothing else. The result is a chosen candidate. The theorem states that if a candidate is chosen among three or more, he or she would not necessarily be chosen in a two-candidate race with each other candidate. If all voters and all candidates are treated alike, the voting method reduces to majority voting in a two-candidate race, and the theorem then states that among three candidates, it is possible that A would get a majority against B, B against C, and yet C against A.

To return to my personal history of this theorem, I will have to start by disregarding the (relatively few) previous works; with the exception of the paper by Duncan Black, to be mentioned below, I knew none of them. As it happens, during my college years, I was fascinated by mathematical logic, a subject I read on my own until, by a curious set of chances, the great Polish logician, Alfred Tarski, taught one year at The City College (in New York), where I was a senior. He chose to give a course on the calculus of relations, and I was introduced to such topics as transitivity and orderings. When, as a graduate student, I came to study mathematical economics with Harold Hotelling, I learned the then not widespread idea that consumers choose commodity vectors as a most preferred point (the ordinalist interpretation), as contrasted with the older view that there was a numerical-valued utility function which they maximized (the cardinalist interpretation). The identity of this view with the logical concept of an ordering was obvious enough. Further, Hotelling and others had already suggested that political choice could follow similar principles of rationality.

My first serious encounter with the idea of social choice came as a curious by-product of economic theory. The standard theory of the firm was based on the reasonable hypothesis that it seeks to maximize profits. Firms however produce over time; for example, investments are made for the purpose of securing production in the future. Since future prices are not known today, the firm cannot really maximize profits over time but only *expected* profits, based on expectations of future prices (and other conditions). In the course of trying to write a dissertation (subsequently abandoned), I was concerned with the fact that firms in the modern world typically had many owners (shareholders). If one ignored the time dimension, this posed no problem; each owner was interested in maximizing profits, and therefore they would all make the same choice. In the more general temporal situation, each owner would want to maximize expected profits. But the owners might easily hold different expectations. Therefore, they would not agree what investment policy would be optimal. The natural assumption was that one investment policy would be chosen over another if a majority of the stockholders (weighted by number of shares) preferred the first to the second. My training led me to check whether this relation is transitive; there was no difficulty in constructing a counterexample.

This struck me as very interesting, but more as an obstacle to my work than anything else. I could not believe it was original and indeed had the conviction that I had heard of this paradox before. I have never been able to verify that I did read it somewhere; it was in fact old, having been developed by the French political philosopher and probability theorist, Condorcet, in 1785. But I certainly had never read Condorcet or any other source I now know of.

I did not pursue the matter any further at that time (1947). The next year, at the Cowles Commission for Research in

Economics at Chicago, I was idly thinking about electoral politics for some reason. I started with a model in which candidates (or parties) arranged on a left-right scale, so that the supporters of any one party would prefer a party nearer to them to one further in the same direction. Thus, each voter had a preference order over parties, with a peak and decreasing in both directions. It was quickly apparent that under these conditions, majority voting would indeed define an ordering. However, about a month after these reflections, I found the identical idea in a paper by Duncan Black in the *Journal of Political Economy*. Again I did not pursue the matter, partly because I felt all this was distracting me from what I took to be my obligation to serious economic work, specifically using general equilibrium theory to develop a workable model as a basis for econometric analysis.

I spent the summer of 1949 as a consultant at the Rand Corporation. This organization had been set up to consult with the United States Air Force. The advice was very freewheeling indeed at that time and included research on the development of the then new concept of game theory. There was a philosopher on the staff, named Olaf Helmer, whom I had met earlier through Tarski; he had translated a textbook of Tarski's for which I read the proofs. He was troubled about the application of game theory when the players were interpreted as nations. The meaning of utility or preference for an individual was clear enough, but what was meant by that for a collectivity of individuals? I assured him that economists had thought about the problem in connection with the choice of economic policies and that the appropriate formalism had been developed by Abram Bergson in a paper in 1938; it was a function, called by him the social welfare function, which mapped the vector of utilities of the individuals into a utility. As stated, this seemed to depend on a cardinal concept of utility, but it seemed to be the case that one could restate it in a manner consistent with the ordinal concept, that is, not a numerically valued utility for individuals and for society but preference orders for both. He asked me to write up an exposition.

I started to do so and immediately realized that the problem was the same I had already encountered. I knew already that majority voting would not aggregate to a social ordering but assumed that there must be alternatives. A few days of trying them made me suspect there was an impossibility result, and I found one very shortly. A few weeks later I made a further strengthening, and this was the form expressed above. A first paper was published in 1950, and a monograph expressing the mathematics more completely and adding a number of interpretive comments in 1951. When presenting the paper at the 1949 meeting of the Econometric Society, a political science professor called my attention to a paper by E. J. Nanson in the *Proceedings of the Royal Society of Victoria* in 1882; in 1952, I learned of the work of Condorcet. Duncan Black has since given a more complete history of the subject. It was marked by rediscoveries, since there had never been a continuous tradition.

NOTES ON CONTRIBUTORS

KENNETH J. ARROW is a Nobel Laureate in Economic Sciences and Professor Emeritus at Stanford University. He was awarded the Nobel Prize in 1972 "for pioneering contributions to general economic equilibrium theory and welfare theory." Arrow began his graduate study in economics and statistics at Columbia University. He has held positions on the Cowles Commission for Research in Economics and at the University of Chicago, Harvard University, and Stanford University. His research, apart from social choice theory, has focused on general economic equilibrium. The profound transformation of the general equilibrium theory is marked by his groundbreaking work. He helped open new productive paths for research in this area and, in so doing, has made fundamental contributions to the renewal of the theory. Kenneth Arrow earned his Ph.D. at Columbia University.

SIR PARTHA DASGUPTA is Frank Ramsey Emeritus Professor of Economics at the University of Cambridge and fellow of St John's College, Cambridge. He is a fellow of the Econometric Society, a fellow of the British Academy, fellow of the Royal Society, foreign member of the American Academy of Arts and Sciences and the American Philosophical Society, honorary fellow of the London School of Economics and of Trinity College, Cambridge, honorary member of the

American Economic Association, member of the Pontifical Academy of Social Sciences, foreign associate of the U.S. National Academy of Sciences, and fellow of the Third World Academy of Sciences. Dasgupta was named Knight Bachelor by Her Majesty Queen Elizabeth II in 2002 for services to economics. Dasgupta's research interests have covered welfare and development economics, the economics of technological change, population, environmental and resource economics, the theory of games, the economics of undernutrition, and the economics of social capital.

ERIC MASKIN is Adams University Professor at Harvard. He received the 2007 Nobel Memorial Prize in Economics (with L. Hurwicz and R. Myerson) for laying the foundations of mechanism design theory. He also has made contributions to game theory, contract theory, social choice theory, political economy, and other areas of economics. He received his Ph.D from Harvard. He was a faculty member at MIT from 1977 to 1984, Harvard from 1985 to 2000, and the Institute for Advanced Study from 2000 to 2011. He rejoined the Harvard faculty in 2012. He is a member of the National Academy of Sciences and a fellow of the Econometric Society, the American Academy of Arts and Sciences, the British Academy, the European Economic Association, the Royal Spanish Academy of Economics and Finance, and the Society for the Advancement of Economic Theory. He is an honorary fellow of St John's College, Cambridge and Jesus College, Cambridge. Among other prizes, he has received the Kempe Award in environmental economics (with P. Dasgupta), the EFR-Business Week Award, the Harvard Centennial Medal, the Cristóbal Gabarrón Foundation International Award in Economics, the Louise Blouin Foundation Award, and the Jean-Jacques Laffont Prize.

PRASANTA K. PATTANAIK is Emeritus Professor in the Department of Economics, University of California, Riverside. His current research interests include welfare economics and the theory of social choice, decision theory, the measurement of deprivation and living standards, and development of economics. He is a fellow of the Econometric Society, and, in 2006–2007, served as President of the Society for Social Choice and Welfare. In addition to his many papers in journals, he has published or edited ten books, including most recently *Essays on Individual Decision-Making and Social Welfare* (Oxford University Press, 2009).

AMARTYA SEN is Thomas W. Lamont University Professor and Professor of Economics and Philosophy at Harvard University and was until 2004 the Master of Trinity College, Cambridge. He is also Senior Fellow at the Harvard Society of Fellows. Earlier on he was Professor of Economics at Jadavpur University Calcutta, the Delhi School of Economics, and the London School of Economics, and Drummond Professor of Political Economy at Oxford University. Amartya Sen has served as President of the Econometric Society, the American Economic Association, the Indian Economic Association, and the International Economic Association. Amartya Sen's books include *Choice of Techniques* (1960), *Growth Economics* (1970), *Collective Choice and Social Welfare* (1970), *Choice, Welfare and Measurement* (1982), *Commodities and Capabilities (1987), The Standard of Living* (1987), *Development as Freedom* (1999), *Identity and Violence: The Illusion of Destiny* (2006), *The Idea of Justice* (2009), and (jointly with Jean Dreze) *An Uncertain Glory: India and Its Contradictions* (2013). Amartya Sen's awards include Bharat Ratna (India); Commandeur de la Legion d'Honneur (France); the National Humanities Medal (United States); Ordem do Merito Cientifico (Brazil); Honorary Companion of Honour (United Kingdom); Aztec Eagle (Mexico); Edinburgh Medal (United Kingdom); the George Marshall Award (United States); the Eisenhauer Medal (United States); and the Nobel Prize in Economics.

JOSEPH E. STIGLITZ received his Ph.D. from MIT in 1967, became a full professor at Yale in 1970, and in 1979 was awarded the John Bates Clark Award. He has taught at Princeton, Stanford, and MIT and was the Drummond Professor and a fellow of All Souls College, Oxford. He is now University Professor at Columbia University in New York. He is also the founder and copresident of the Initiative for Policy Dialogue at Columbia. In 2001, he was awarded the Nobel Prize in Economics for his analyses of markets with asymmetric information, and he was a lead author of the 1995 "Report of the Intergovernmental Panel on Climate Change," which shared the 2007 Nobel Peace Prize. In 2011, *Time* named Stiglitz one of the 100 most influential people in the world.