LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

VOLKSWIRTSCHAFTLICHE FAKULTÄT

Claude Hillinger:

On the Possibility of Democracy and Rational Collective Choice

ON THE POSSIBILITY OF DEMOCRACY AND RATIONAL COLLECTIVE CHOICE

Claude Hillinger*

JEL Classification: D71, 72

Key words: Arrow's paradox, approval voting, cardinal collective choice, democracy, instant runoff voting, plurality voting, voting paradoxes

October, 2004

*SEMECON
University of Munich
Ludwigstr. 33/IV
D-80539 Munich
Germany
e-mail: hillinger@econhist.de

ABSTRACT

The paper challenges the 'orthodox doctrine' of collective choice theory according to which Arrow's 'general possibility theorem' precludes rational decision procedures generally and implies that in particular all voting procedures must be flawed. I point out that all voting procedures are cardinal and that Arrow's result, based on preference orderings cannot apply to them. All voting procedures that have been proposed, with the exception of approval voting, involve restrictions on voters expressions of their preferences. These restrictions, not any general impossibility, are the cause of various well known pathologies. In the class of unrestricted voting procedures I favor 'evaluative voting' under which a voter can vote for or against any alternative, or abstain. I give a historical/conceptual analysis of the origins of theorists' aversion to cardinal analysis in collective choice and voting theories.

1. INTRODUCTION

> democracy: Government by the people, exercised either directly or through elected representatives.
> *The American Heritage Dictionary of the English Language*
>
> What Kenneth Arrow proved once and for all is that there cannot possibly be found such an ideal voting scheme: The search of the great minds of recorded history for the perfect democracy, it turns out, is the search for a chimera, for a logical self-contradiction.
> Paul Samuelson (1972) on the occasion of Arrow winning the Nobel Prize.

Samuelson here expresses the *orthodox doctrine* in collective choice theory and more generally in contemporary political science. While this paper aims to refute the orthodox doctrine, I begin by explicating it.

Samuelson's choice of words is somewhat misleading: The "great minds of recorded history" were more interested in the search for a workable democracy than a "perfect" one. As Winston Churchill put it rather drastically: "It has been said that democracy is the worst form of government except all the others that have been tried." Arrow (1951) did not claim to have discovered a minor blemish on the surface of an otherwise perfect democracy concept. His claim is that no collective choice procedure can satisfy a set of requirements that appear very basic and elementary[1]. If we define as 'reasonable' a collective choice procedure that satisfies Arrow's criteria, then no reasonable collective choice procedure exists and if that is the case, it would appear that 'democracy', by any common definition of the word cannot exist either.

The orthodox doctrine views Arrow's result as perfectly general and thus applicable to any procedure for collective choice, in particular to voting. The history of voting theory appears to confirm this view. Formal voting theory began with the work of Borda and Condorcet[2]. Both authors were motivated by the desire to find a voting method that would improve on plurality voting (PV). At present, 200 years later, a immense literature on voting has been produced. Theorists agree with Borda and Condorcet on the flawed nature of PV, but there is no agreement on a superior voting rule. Indeed, based on Arrow's result, there seems to be little hope among theorists that such a superior method can be devised. In their introduction to a recent symposium on voting theory, Levin and Nalebuff (1995, p.3) put it as follows:

---

[1] Since Arrow's postulates are discussed in every introductory textbook on collective choice, as well as in a vast journal literature, I omit stating or discussing them here.

[2] On early voting theory see McLean and Urken (1995).

One can speculate on why alternatives to plurality rule have had such a difficult time being adopted. Part of the cause may be Arrow's general possibility theorem. Arrow (1951) demonstrates that any voting system applied to an unrestricted collection of voter preferences must have some serious defects; we must always choose between flawed alternatives.

My challenge of the orthodox doctrine can be summarized in the following assertions:

I: Arrow's theorem proves the impossibility of meaningful aggregation on the basis of orderings.

II: All voting methods involve cardinal aggregation and are therefore not subject to Arrow's result.

III: Common voting methods restrict the expression of voter preferences. When these restrictions are lifted, satisfactory voting methods result.

The theories of collective choice and of voting have been primarily concerned with so-called 'paradoxes', meaning sets of reasonable appearing assumptions, not all of which can be satisfied. If the above assertions are true, a more fundamental 'paradox' is the adherence of theorists to the orthodox doctrine. In Section (6) I try to find the origin of this *paradox of voting theory*.

As my entire approach differs from that of the orthodox doctrine, I require some different concepts and terminology. I define these at the outset and elaborate as needed later on. I conceptualize voting as a problem of *measurement*. If preferences are to be aggregated, they must first be measured. Measurement requires a *scale* that attaches numbers to whatever it is that is being measured. I will discuss various *voting scales* that are used in different voting methods. A voter's preference among the alternatives is measured by the numbers, called *scores*, that he attaches (implicitly or explicitly) to each alternative. All election methods determined the outcome by adding the scores and are therefore *cardinal*.

In welfare economics and social choice theory a social welfare function (SWF) is written as $W(u_1,...,u_N)$, where the u_i, are the relevant individual, or more problematically household, utilities. If W takes the form $W = a + \sum b_i u_i, \quad b_i > 0$, or more simply $W = \sum u_i$, it is referred to equivalently as *cardinal*, or *utilitarian*. In the general case it is usually assumed that the u_i are defined up to a positive, monotone transformation. For the utilitarian SWF the usual assumption is that individual utility is of the von Neumann – Morgenstern type, i.e., defined up to a positive, linear transformation. A further assumption that is universally, usually implicitly, made must be made explicit in the context of the present paper. It is the assumption that the individual is free in his assignment of utility values to the alternatives that confront him. In economics this assumption is linked to the concept of *consumer sovereignty*. In a

voting context I refer to it as *voter sovereignty*. The reason for emphasizing these distinctions is that all common voting methods, with the exception of approval voting (AV), violate consumer sovereignty. I refer to these as *restricted voting* (RV) methods.

*A*ll those voting methods which allow voter to freely assign to each alternative the scores made available by a given voting scale will be termed *unrestricted voting* (UV) methods. Specifically, AV is a UV method with a 2-valued voting scale, or UV-2. Evaluative voting (EV), proposed in this paper is UV with a 3-valued scale, i.e., UV-3.

Another important issue in relation to SWFs is how 'utility' is to be interpreted and measured. This is usually left open by the advocates of the SWF approach. As described in Section (6), when the matter is discussed, it appears that 'utility' is viewed as a comprehensive measure of psycho-physical well being. I believe that such a concept may be appropriate in psychology or medicine, but is inappropriate in economics and politics, where I am unaware of a single operationalization of such a concept. Instead, I propose the term *utility indicator (UI)*, for any measure that effects individuals', or families', wellbeing in a concrete situation and, importantly, is sufficient to determine the individual's choice in that situation. A function that aggregates utility indicators over a group, so as to obtain a criterion for a group decisions, will be termed a welfare *indicator* (WI).

In a series of papers on economic welfare measures (Hillinger, 2001, 2002, 2003a, 2003b), I argued that the change in real expenditure (RE), defined as money expenditure at constant prices, is such an indicator at both household and social levels. By the nature of expenditure, if RE_i is a UI for the *i*th individual, $RE = \sum RE_i$ is a WI. Particular cases are: aggregate consumer surplus, $CS = \sum \frac{1}{2}\left(p_j^0 + p_j^1\right)\left(q_j^1 - q_j^0\right)$, or real national product, $RNP^t = \sum p_j^0 q_j^t$, where the quantities are defined at the appropriate level of aggregation and the market prices are assumed independently given. All of these measures are particular approximations to RE.

Turning to voting, let s_{ij} be the score of the *i*th voter for the *j*th alternative and $s_j = \sum_i s_{ij}$ the total score of the *j*th alternative. The aggregator function $S(s_{1j},...,s_{Nj}) = s_j$ is a WI and the alternative that maximizes it is the social choice. Multistage voting procedures are not materially different.

The paper has the following structure: Section 2 describes the most important voting methods in terms of the voting scale used and the restrictions imposed on it. Section3 defines the general class of *evaluation scales* that are in use, particularly with regard to how many divisions they have. I conclude the EV, with a 3-valued scale is appropriate for general

elections. The overall argument for EV is summarized in Section 4. The defects of PV and related voting rules are discussed in Section 5. A novelty here is that the problem of *tyranny of the majority*, usually treated in constitutional theory is related to voting theory. Under the heading *paradox of voting theory*, Section 6 discussed the self-defeating fixation of voting theorists on the ordinal approach. Section 7 concludes the paper.

2. THE DESCRIPTION OF VOTING METHODS

2.1: The Framework

The purpose of this section is to make the description of voting methods given in the Introduction precise and apply it to the most important ones. It follows from this description that all voting methods satisfy two postulates regarding voting scales and cardinal aggregation. When a third postulate on voter sovereignty is added, UV results.

Definition: Voting Scales

a. A voting scale is a set of numbers, $S_k = (s_k, s_{k-1}, ..., s_1)$, called scores, such that $s_j = s_{j-1} + d, \quad d > 0$.

b. Given a voting scale S_k, we can define another valid voting scale $S'_k = (s'_k, s'_{k-1}, ..., s'_1)$, with $s'_j = a + bs_j, \quad b > 0$.

In the following I will assume $d = 1$, which corresponds to the voting scales in actual use.

Crucial features of voting scales are that they are *truncated* and that the scale is the same for all voters under a given voting method. This assures that all voters have the same *power* to influence the outcome.

2.2 Empirical Voting Scales

The following examples show that all voting methods satisfy two postulates:

Postulate VS: A voting method must be based on a voting scale as defined above.

Postulate CA: The outcome of an election must be based on cardinal aggregation ,i.e., the simple sums of scores for the various alternatives.

Since the decision is always arrived at by summing the individual scores, I don't state this explicitly for each voting rule. In the following I limit myself to a discussion of voting rules that have been prominent in the literature, or in practice, or that are relevant from the perspective of the present paper. I begin the discussion with the Borda count, because, though

it is cardinal like all voting procedures, it corresponds most closely to the idea of an ordinal scale.

a. Borda count (BC)

The voting scale is defined by $s_1 = 1$, k equal to the number of alternatives. The restriction is that each score must be used once.

b. Plurality voting (PV)

The scale consists of the two values (1,0). The restriction is that 1 can be given to only one alternative, with the rest receiving 0.

c. Majority rule (MR)

This is the special case of PV when there are just two alternatives.

d. Cumulative voting (CV)

In this method, also known as bucket voting, each voter is given a fixed number of votes, usually equal to the number of alternatives, and is free to distribute these over the alternatives in any manner she chooses. The scale is therefore the same as for the Borda count, and the restriction is given by the requirement $\sum s_j = k$.

Multi-stage voting rules

e. Instant Runoff Voting (IRV)

This is a variant of the method called *single transferable vote* (STV), or Hare system, involving a computer simulation of a hypothetical multistage voting process, on the basis of information gathered in the first round. I describe it instead in terms of the hypothesized multistage process.

Each round is identical with the PV process, the difference being that when there is no majority winner, the alternative with the lowest score is removed before the next round.

f. Condorcet rule (CR)

To win an election a candidate must defeat each opponent separately by MR. Each stage of the process is thus MR.

Unrestricted Voting (UV)

g. Approval voting (AV)

As in plurality voting, there is a (1,0) scale. The plurality rule's restriction on the allocation of these values to the alternatives is lifted so that either value can be given to any alternative. AV is UV-2.

h. Evaluative voting (EV)

The scale is (1,0,-1) and there is no restriction on the assignment of scores. EV is therefore UV-3.

Postulate VS: A voting rule must allow for voter sovereignty, meaning that the voter must be free to assign to any alternative any of the scores provided by the voting scale.

All methods discussed satisfy Postulates CS and CA. The final two satisfy in addition Postulate VS and are therefore members of the class UV.

Nothing has thus far been said about the choice of k, the number of alternative values provided by a voting scale. This is a topic of the next section.

3. A SURVEY OF EVALUATION SCALES[1]

Voters express their electoral preferences by evaluating the alternatives. The evaluations are transformed into numerical scores and are subsequently added to obtain the election result. The sequence a: evaluation, often by means of descriptive adjectives, b: assignment of numerical scores, c: collective evaluation by addition of the scores, is employed in many other contexts. Examples are: Surveys of customer satisfaction conducted by firms and other organizations. Political surveys of voter attitudes towards issues or candidates. The scoring of athletic performances in contests. Finally, the grading of students in schools and universities. There is no unified field of study or terminology regarding the examples cited. Some are dealt with in academic social science under the headings of 'attitude scales', or 'opinion research'. Of interest in the context of the present paper are primarily the scales employed. I use the general tem '*evaluation scale*' to refer to them. A voting scale is an instance of an evaluation scale.

The scales have in common that they measure, in different proportions, elements of preference and of cognition. At one extreme, a restaurant customer asked to rate her dinner, does so on the basis of her preferences, i.e., her like or dislike of what she just experienced. At the other extreme, a teacher grading an exam is expected to perform a purely cognitive task

---

[1] A fascinating history of social, to some extent also of physical, measurement is Duncan (1984). Interestingly, the first chapter deals with the history of voting, regarded by Duncan, as by the present author, as an instance of social measurement. The scales discussed in this paper are referred to in social science research as *rating scales,* and are a special case of *attitude scales.* A good survey of these scales and their history is Dawes (1972).

based on her professional training. She is expected not to let her preferences, for example her personal likes and dislikes regarding the students, have an effect on the score.

When scoring candidates in a general election, both preferences and cognition play major roles. A voter's ultimate preference may be for a prosperous and peaceful home country and if possible world. Which candidate, once elected, will best serve these objectives will be highly uncertain. Candidates may not be very specific about their plans and their later actions often bear little relationship to their earlier pronouncements. Even when a candidate does as promised, the consequence may be very different from what she and her supporters expected.

Irrespective of the relative weights of preference and cognition, all scoring procedures satisfy postulates VC and CA. They are also quite uniform regarding the number of divisions of the scale. A common type of questionnaire ask for some service,to be rated on a scale of (excellent, good, satisfactory, average, poor, unsatisfactory), of course, equivalent different terms may be used. To obtain an average rating. these terms are translated to a scoring scale such as (5,4,3,2,1), or (2,1,0,-1,-2) and the sum is divided by the number of responses. Similarly, in final grades for students, it is universal practice to use a scale of five, maximally 6 different scores that range in interpretation from 'excellent', or 'exceptional', to 'failed'.

The number of divisions of a scale is a pragmatic issue related both to the ability to discriminate and to the possible gain from an increase in precision. I advocate the EV-scale (1,0,-1). The motivation for this choice is: *a*. Voters are typically ill informed and faced with a high degree of uncertainty regarding the ultimate consequences of their choices. There would be little genuine informational gain in using a 5-point, rather than a 3-point scale. *b*. When it comes to specifying a voting scale, there is also the difficult and important problem of convincing voters and politicians to accept it. I believe that EV, which allows each voter to rate each candidate positively, or negatively, is intuitively the most attractive voting rule. *c*. Relative to PV, EV offers a vast extension of choice possibilities. as is illustrated by Table 3.1 In case of 5 alternatives, AV offers 6 times as many choices as PV, EV 48 times as many. The cases for which all candidates receive the same score, equivalent to abstention, were removed.

Table 3.1: Number of Possible Votes

Alternatives	PV	AV	EV
N	N	$2^N - 2$	$3^N - 3$
5	5	30	240

4. EVALUATIVE VOTING: A SUMMARY OF THE ARGUMENTS

In this section I summarize the arguments for unrestricted cardinal voting in general and for EV in particular.

I. Voter sovereignty

There is no justification for restricting voters' freedom in expressing their preferences of a given scale. In analogy to the economic concept of 'consumer sovereignty', I refer to this principle as '*voter sovereignty*'. Among common voting methods, only EV respects voter sovereignty. Voter sovereignty is closely related to Arrow's condition of unrestricted domain. The violation of either condition implies that certain patterns of preference cannot be expressed by the voter. Imposing voter sovereignty on any current voting system implies UV with the scale of the given voting rule.

II. The universal practice of measurement

All of the principal scales that we use to measure the physical universe are cardinal. Ordinal scales, such as the Richter earthquake scale, are few and of little theoretical importance since they are not in a causal relationship with other measured magnitudes. A similar statement applies to psychological phenomena that I have jointly characterized as 'evaluations', a voter's evaluation of the alternatives are one example. It is only in voting, and particularly in political voting, that the scales are restricted. For this there is no apparent reason, nor, as far as I know, has any argument in defense of this practice been advanced.

III. The argument for EV

The scale used in any particular measuring context is a pragmatic issue. It is related both to the achievable accuracy and its cost, as well as to the accuracy needed in the given application. Paramount in the voting context should be that the scale is regarded as appropriate by a majority of voters. I have not been able to make a systematic and large scale study, but in personal conversations, those to whom I explained EV have found it very attractive. I am personally convinced that given the chance to choose, most voters would prefer EV. Voters generally have scant information by which to evaluate the candidates, they do know whom they are for, whom they are against and about whom they are indifferent or know too little to have an opinion.

Thus far I have made the positive case for EV. The choice of any one alternative depends, of course, on the other alternatives available. These are discussed in the next section

5. THE PROBLEM WITH PLURALITY VOTING AND RELATED SYSTEMS

That an unpopular candidate may win under PV is well known and has provided the major impetus to the search for alternatives. Less well known is the fact that the most popular candidate may come in last and may therefore be eliminated in a multistage election procedure. This is the focus of the following sections.

5.1 Plurality Voting

For those who have studied voting mechanisms in the hope of improving them, the principal motivation has been to cure the defect of PV, that it may lead to the election of a minority candidate. Various examples of such pathological outcomes are given by Brams and Fishburn (1983). Tabarrok and Spector (1999) argue that Lincoln was a minority candidate.

My interest in the subject was aroused by thinking about the US presidential elections in 1952 and 1956. The Democratic candidate in both elections was Adlai Stevenson, an intellectual with a sharp self-deprecating wit, the darling of the college crowd and rather disliked by the American mainstream. In the primaries Stevenson was opposed by bland mainstream candidates. Stevenson was elected because the mainstream vote was divided among his opponents. No Democratic candidate could have won against 'Ike' in these elections, but the fact remains that the Democrats probably chose their weakest candidate.

PV elections that result in undesirable outcomes are of two types: the first is illustrated by US presidential primaries with many candidates, the second by US presidential elections when there is a third party candidate. A recent example is the candidacy of Ralph Nader in the presidential election of 2000. Most Nader supporters would probably have voted for Gore rather than Bush. In this very close and contested election, in which Gore actually received more popular votes than Bush, Nader's candidacy probably tilted the scale in Bush's favor. It is typical of such elections that the third party candidate takes votes away from the candidate that he is ideologically closest to and perversely favors the election of the candidate that he and his supporters most oppose.

From a more formal point of view, the problem with PV is that only the voters' most favored alternatives are considered. A possible consequence is that an alternative that is the worst for a majority of voters may actually be chosen. Not only is it possible that the most popular candidate will not be chosen, he may receive the least number of votes. As explained in the next subsection, this fact is crucial in evaluating the system of instant runoff voting (IRV) that is currently gaining popularity in the US. A simple example is given in Table 5.1, where *abc* means *a* is preferred to *b* is preferred to *c*.

Table 5.1: Most Disliked Candidate Wins Under PV, the Most Popular Loses

Preference	*abc*	*cba*	*bca*
No. of Voters	40	30	20
EV Score	1,0,-1	-1,0,1	-1,1,0

In this example, b is the Condorcet winner. The results of pair wise contests would be: *ba*(50/40), *ca*(50/40), *bc*(60/30). Candidate *a*, who would loose every pair wise contest, wins the PV election. The Condorcet winner *b* comes in last. The final row gives the EV scores of voters of the given type, always in the order (*a,b,c*). The aggregate score is (-10,20,-10). Candidate *b* is the hands down winner under EV.

5.2 Multistage Voting Systems

The most widely used two-stage voting system is PV combined with a runoff vote between the two candidates with the most first place votes. A multistage voting system that has some popularity in English speaking countries is the single transferable vote (STV), also known as Hare system. The STV was originally intended for elections in which several members of an assembly are to be chosen from one district. I discuss here only the case of the election of a single candidate. Voters are asked to score the candidates according to the Borda rule. If there is a candidate with a majority of first place scores, he wins. If not, the candidate with the fewest first place scores is eliminated. A second round of voting than takes place. The process continues until a candidate wins a majority of first place scores. At the latest, this occurs when there are only two candidates left.

A variant that is being advocated with considerable success in the United States, by an organization named Center for Voting and Democracy, (CVD) is instant runoff voting (IRV). The principal innovation of IRV is that computer computations are substituted for repeated rounds of voting. When there is no first round winner, the computer eliminates the candidate with the fewest first place scores and advances the second place choices of those who voted for him to first place. The new first place choices are added by the computer to see if there is a winner of a majority of first choices. If not, the process is repeated, again by computer computation.

A defect of multistage systems that has long been known is that they lack monotonicity. By this is meant that an increase of support to an otherwise winning candidate may lead to his elimination. This may occur in two different ways:

The No-Show Paradox

By voting their sincere preferences a group of voters may turn their first choice from a winner to a looser.

The Additional Support Paradox

When a candidate moves up in the preferences rankings of some voters, without moving down in the ranking of any voter, he may change from winner to loser.

Nurmi (1999) discusses monotonicity paradoxes in detail and shows that they apply to most multistage voting systems.

The advocates of STV, or IRV, argue correctly that PV has the defect of not taking the second, or third, place preferences of voters into account. The claim that STV, or IRV do take this into account is false. The elimination at each round depends on first place choices only. If, as in the above example, the most popular candidate on the basis of all scores is eliminated in the first round, the consideration of some second pace ratings in subsequent rounds cannot remedy this defect.

Another defect of IRV is that there is no way of verifying the complex computations that the computer must perform. The usual recount of a doubtful election becomes impossible. This is a general problem of electronic voting, but it is exacerbated by IRV.

On the web page of the CVD I could not find any argument in favor of IRV that had intellectual or scientific substance. In particular, they do not refer to the well established defects of their method. The organization is however an advocacy powerhouse, led by for congressman and independent presidential candidate John B. Anderson. They boast that IRV has become much more popular than AV, that has long been advocated by political scientists Brams, Fishburn and others. I find it disturbing that, on an issue so vital to the functioning of democracy, snake oil sells better than a genuine remedy.

5.3 Strategic Voting

Thus far I assumed that voters are guided solely by their preferences in determining their vote. This is the assumption of sincere voting. A major criticism of PV as well as of the Borda rule used in both STV and IRV is that these voting rules provide a strong incentive to vote strategically. Assume that the candidate most favored by a voter has little chance of being elected. Since a voter cannot give his best score to more than one candidate, he has an incentive to give it to the most preferred among those having a realistic chance. By voting in this strategic manner, the voter reveals a false preference that is not his true preference. Under

utilitarian voting each candidate is rated independently and the incentive to strategic voting is reduced.

6. THE PARADOX OF VOTING THEORY

6.1 Introduction

For both theoretical and empirical reasons, one might have expected social choice theorists to abandon the assumption of ordinal preferences in favor of a cardinal theory. Arrow showed that as long as only ordinal rankings are admitted, no reasonable method of collective choice exists. Starting with Fleming (1952) and Harsanyi (1955)[1], many authors have given plausible, self consistent derivations of cardinal social welfare functions , i.e., functions of the form $W = \sum u_i$, where W is social welfare and the u_i are the utilities of the society's members. Sen (1986, p. 1073) accurately described the resulting state of collective choice theory:

> some conditions that look mild – and are indeed satisfied comfortably by utilitarianism when translated into its cardinal interpersonally comparable framework – cannot be fulfilled by *any* rule whatsoever that has to base the social ordering on n-tuples of individual orderings.

The empirical fact is that all voting procedures assign numerical values to the expressed preferences of voters and add these to obtain a result. When evaluation scales are used in such fields as psychology, sociology, or survey research, this is explicitly recognized and incorporated into the relevant theories. The theories of collective choice and voting appear to be the only exceptions.

Both the theoretical and the empirical evidence for the need of a cardinal theory are overwhelming. Given this evidence, one would have expected theorists to reject ordinalism in favor of cardinalism. That this did not happen is the paradox of voting theory. In this section I try to shed some light on its origins.

Sen (1986) begins his review of social choice theory by noting than it arose from the conjunction of two initially separate streams of theorizing, each reaching back about 200 years. One is the tradition of social welfare analysis based on individual wellbeing, that in its utilitarian form goes back to Jeremy Bentham and the classical economists and eventually took the form of modern welfare economics. The other source, mathematical voting theory,

---

[1] A good exposition of his mature view on cardinal social welfare is Ch. 4 in Harsanyi (1977).

can be traced back to Borda and Condorcet. Regarding the two traditions Sen (1986, p.1073)writes:

> No approach to welfare economics has received as much support over the years as utilitarianism....It is clear that utilitarianism uses cardinality and interpersonal comparability of personal utilities. Both these practices received severe reprimands in the 1930's, with the rebuke drawing sustenance from a single minded concern with basing utility information on non-verbal behavior only, dealing with choices in the absence of risk. It thus appeared that social welfare must be based on just the n-tuple of ordinal, interpersonally non-comparable, individual utilities…
>
> The other source, dealing primarily with election methods, had in any case the tradition of concentrating on the information given by an n-tuple of individual orderings – reliant on an informational framework that was much less ambitious than utilitarianism…
>
> The union produced modern social choice theory. The big bang that characterized the beginning took the form of…Arrow's "General Possibility Theorem".

Sen's discussion raises a central foundational issues of collective choice theory: Why did welfare economics abandon its long cardinal tradition? Why did voting theory commit to ordinalism? Sen does not attempt to answer these questions. In searching for answers, I will deal first with welfare economics and then with voting theory.

6.2 Welfare economics

The attack on cardinalism began with the marginalists' assault on classical economics and on its philosophical background, which was utilitarianism. The marginalists claimed that utility was not measurable and that therefore classical economics was metaphysics more than science. This can be regarded as the kind of hyperbole that is not unusual when one school of thought challenges another. What is more surprising is that the issue of ordinalism vs. cardinalism gained new actuality in the 1930's when the battles between marginalists and the classical school were history. Lionel Robbins' *Essay on the Nature and Significance of Economic Science* (Third ed. 1984, first published in 1932), that vigorously championed the ordinalist position, became a kind of bible on methodology for mainstream economics. It replaced in that function John Stuart Mill's utilitarianism as the standard account of the utilitarian position in economics.

I believe that the timing and enormous success of Robbins' treatise is explained by its role in the struggle between socialism, than at the peak of its power and influence, and economic orthodoxy. Explicitly, Robbins argued that what remains of economics when interpersonal utility comparisons are renounced is a value free science. Implicitly, he pulled the intellectual rug from under the interventionist and usually redistributive policy proposals of the socialists.

The answer to Robbins and the new economic orthodoxy was not long in the coming. It was formulated by the most distinguished economist of the left and Nobel Prize winner Gunnard Myrdal. In several publications he argued that all science has an irreducible value

component and that the denial of this leads to the pursuit of hidden agendas. His solution was to urge all scientists, including economists, to state their values explicitly.[1]

Unfortunately in my view, Myrdal's criticism had no impact on mainstream welfare economics. There the ordinalist position was merged with the idea that interpersonal utility comparisons are impossible. The originator of the "big bang" that started modern social choice theory states a clear position:

> The viewpoint will be taken here that interpersonal comparison of utilities has no meaning and, in fact, that there is no meaning relevant to welfare comparisons in the measurability of individual utility. The controversy is well-known and hardly need be recited here. During the entire controversy, the proponents of measurable utility have been unable to produce any proposition of economic behavior which could be explained by their hypothesis and not by those of the indifference-curve theorists.1 Indeed, the only meaning the concepts of utility can be said to have is their indications of actual behavior, and, if any course of behavior can be explained by a given utility function; it has been amply demonstrated that such a course of behavior can be equally well explained by any other utility function which is a strictly increasing function of the first. If we cannot have measurable utility, in this sense, we cannot have interpersonal comparability of utilities a fortiori. Arrow (1951, 1963, p.9).

These remarks are followed by a discussion of von Neumann – Morgenstern utility, asserted to be irrelevant for social choice.

> Even if, for some reason, we should admit the measurability of utility for an individual, there still-remains the question of aggregating the individual utilities. At best, it is contended that, for an individual, his utility function is uniquely determined up to a linear transformation; we must still choose one out of the infinite family of indicators to represent the individual, and the values of the aggregate (say a sum) are dependent on how the choice is made for each individual. In general, there seems to be no method intrinsic to utility measurement which will make the choices compatible.(Footnote) It requires a definite value judgment not derivable from individual sensations to make the utilities of different individuals dimensionally compatible and still a further value judgment to aggregate them according to any particular mathematical formula. If we look away from the mathematical aspects of the matter, it seems to make no sense to add the utility of one individual, a psychic magnitude in his mind, with the utility of another individual. Even Bentham had his doubts on this point.
>
> We will therefore assume throughout this book that the behavior of an individual in making choices is describable by means of a preference scale without any cardinal significance, either individual or interpersonal.

I cite the footnote also since it illuminates most clearly the fundamental issue involved:

> It must be granted, though, that, if it is assumed to begin with that all preference scales for individuals are the same (all individuals have the same tastes), then we could choose the utility function the same for all. However, if we take seriously the idea of interpersonal comparison of utilities, we must allow for the possibility that, of two individuals with the same indifference map, one is twice as sensitive as the other, and so the proper utility function for one should be just double that for another. It would be interesting, indeed, to see an operational significance attached to this concept of differing sensitivity.

In the vast literature that was inspired by Arrow's book, the above statements, as well as the ordinal approach taken, have hardly been questioned. They may therefore be taken as

---

[1] Myrdal's thought is discussed in Streeten (1987).

representative of the orthodox position in modern collective choice theory. Before criticizing this view, I state what I take to be its most essential features:

a. Utility is a psycho-physical measure of the intensity of feelings

This is a clear implication of the passages cited above. Even more explicit is the following passage from a recent article by Tullock(2004):

> Let me, however, introduced a little science fiction. We do not know now what will be discovered in the next hundred years. As we learn more about the brain it may be possible for us, with the aid of supercomputers, to determine how strong various people's feelings are. Probably the rating of the alternatives in the Borda method is to some extent an indication of the strength of people's feelings on the various alternatives. It is of course a very poor measure, but there's no reason to believe that we may not have, lOO years in the future, a very good measure.

> If we had such a measure we might prefer not that alternative, which has the highest number of first preferences, but that alternative which maximizes the total satisfaction I received by the voters. Thus if Mr. 5 had only a rather weak preference of alternative A I over alternative B, we might wish to discount that preference somewhat. Maximizing total satisfaction, if we could do it, would seem to be better than simply selecting the alternative, which is preferred by the most voters even if in some cases that preference is very weak.

In other words, Tullock would be a cardinalist if he thought that a psycho-physical measure of the intensity of feelings could be obtained. Actually, he is underestimating the speed of technological progress. Political consultants already use brain scans to measure the emotional responses to political advertisements. It would be technologically feasible in the 2004 presidential election in the US to present voters with pictures of Bush and Kerry, compute the responses via brain scans and add the results to determine the election outcome. Should we do this? I argue below that we should not.

b. Utility is an abstract philosophical concept with an implicit claim of universal applicability.

Arrow's general possibility theorem[1] is, as the name implies, perfectly general. By virtue of its negative character, it follows a fortiori that no solution to any problem is claimed. Harsanyi and others offered non-contradictory axioms for cardinal choice. They do not, however, state how these utilities are to be measured in any concrete instance. For the Bergson – Samuelson SWF, the indeterminacy is even greater; not only the measurement of utility, but also the form of the SWF are left open.

As discussed in the introduction, the SWF approach was unable to connect with the quantitative measures of applied welfare economics. In order to provide a theoretical foundation for these measures I proposed the concepts of context based utility indicators (UI)

---

[1] This is Arrow's own, rather confusing terminology. More often the result is referred to as 'Arrow's impossibility theorem', or 'Arrow's paradox'.

and welfare indicators (WI) and I argued that individual and aggregate voting scores are such indicators. Both in welfare economics and in voting theory, if theory is to have contact with reality it must move from the perfectly general, and necessarily undefined, to the more specific. A part of this must be the movement from ordinalism to cardinalism.

6.3 Voting Theory

The only attempt an interpretive history of voting theory of which I am aware is the introductory chapter in McLean and Urken (1995), who cover a period from antiquity to Nanson's contribution in 1882. They stress the discontinuity of voting theory which was repeatedly started and then lapsed, with earlier contributions being forgotten. Miller (1987), somewhat unfairly, writes:

> …Historically, economists have contributed at least as much as political scientists to the pure theory of voting. The theory of voting has its origins in the work of such enlightenment philosophers and mathematicians as Borda, Condorcet and Laplace. Little further progress was made until some forty years ago when the economist Duncan Black wrote a series of articles (most notably Black, 1948) on the logic of committees and elections, which were subsequently consolidated into a book (Black, 1958). Since Black revived the subject, a number of economists and political scientists have made important contributions. Indeed, the theory of voting has to some extent been subsumed by the more recent and abstract theory of social choice, which was virtually invented by the economist Kenneth Arrow (1951).

My aim here is to give a plausible explanation of the *paradox of voting theory* defined as: *a*. The prominence of plurality voting, Borda count and Condorcet criterion in voting theory, in spite of their generally understood defects. b. The failure to view the problem as one of measurement, which would have led straight to UV.

The answer I propose is that voting theory can be seen as a series of attempts at generalizing the binary case, number of alternatives T=2, to the case T>2. Different voting methods generalize different aspects of the binary case. PV is the easiest, EV the least obvious of these generalizations. The generalizations involved in the different voting methods will now be examined.

PV: This is the simplest generalization of MR: The voter simply continues to vote for a single alternative.

CC: This is also straightforward: The problem of multiple alternatives is reduced to a series of binary choices, to be decided by MR.

BC: The generalization here is less obvious: In a binary choice the alternatives are rated by means of the (1,0), with the restriction that each value must assigned to a candidate. Borda's proposal extends the scale, so that for T alternatives it becomes (*T,T-1,…,1*), with the same restriction as before.

The various forms of UV, including AV and EV can be derived from MR under a more general interpretation of the MR voting scale that allows (0,0) and (1,1) votes. Voters with these preferences will generally not vote, but there is no problem if they do, since the outcome is not effected.

AV: Under the interpretation just given, any candidate can receive 0, or 1. If this rule is maintained for an arbitrary number of candidates, AV results.

EV: To go from MR to EV requires more imagination than in the cases just discussed. One has to see the (0,1) scale of MR as a genuine *measuring* scale and think further about the optimal number of distinct values for a voting scale. The fact that this route is not intuitive explains why it was not traveled and thus provides an answer to the paradox of voting theory.

In my case, the route started with cardinal welfare economics, made a detour through cardinal measurement in psychology, sociology and opinion research to finally arrive at the idea that voting depends on the cardinal measurement of preferences.

Modern mathematical economics, which has largely shaped collective choice theory, has evolved to give primacy to formal theories and has produced a lack of concern for measurement, or more generally in the processes that generate data. This mental frame has made it particularly difficult to conceive of social choice as a measurement problem

7. CONCLUSION

The paper challenges the 'orthodox doctrine' of collective choice theory according to which Arrow's general possibility theorem precludes rational collective choice and therefore also democracy. I point out the Arrow's result is valid only under the assumption that preferences are expressed orderings, whereas all voting procedures are cardinal, though generally restricting how preferences can be expressed. When the restrictions are lifted, in accordance with the postulate of consumer sovereignty, unrestricted voting results. With a 2-valued scale this is approval voting, with a 3-valued scale evaluative voting.

I advocate evaluative voting that allows a voter with regard to any alternative, to vote for, or against it, or to abstain. Allowing voters to freely express their evaluations, positive or negative, is not only rational, but in addition better meets the emotional needs of voters and is therefore likely to increase participation.

The paper also discusses the 'paradox of voting theory', meaning collective choice theorist fixation on Arrow's negative result in spite of the fact that a number of consistent derivations of cardinal SWFs are available. I hope to develop further the relationship between cardinal SWFs and voting in another paper.

8. REFERENCES

Arrow, K. J. (1951), Social Choice and Individual Values, Wiley, New York. Second edition, 1963, New Haven, Yale University Press.

Brams, S. J. and Fishburn, P. C. (1983), Approval Voting, Boston, Birkhäuser.

Dawes, R. M. (1972), Fundamentals of Attitude Measurement, John Wiley, New York

Duncan, O. D. (1984), Notes on Social Measurement: Historical and Critical, New York, Russell Sage Foundation.

Eatwell , J., Milgate, M., Newman, P. (eds.), (1987), The New Palgrave: A Dictionary of Economics, London, Macmillan Press.

Fleming, M. (1952), A cardinal concept of welfare, Quarterly Journal of Economics, LXVI, 366-84.

Harsanyi, J. C. (1955), Cardinal welfare, individualistic ethics, and the interpersonal comparisons of utility, Journal of Political Economy, LXIII, 309-321. Reprinted in (1976), Essays on Ethics, Social Behavior, and Scientific Explanation, Dordrecht, D. Reidel.
(1977), Rational Behavior and Bargaining Equilibrium in Games and Social Situations, Cambridge University Press.

Hillinger, C. (2001), Money metric, consumer surplus and welfare measurement, *German Economic Review,* 2, 2, 177-193.
(2002), Consistent aggregation and chaining of price and quantity measures. *Journal of Economic and Social Measurement*, 28, 1-20.
(2003a), The money metric, price and quantity aggregation and welfare measurement, Contributions to Macroeconomics, 3, 1, art7. The article can be downloaded at .http://www.bepress.com/bejm/contributions/vol3/iss1/art7.
(2003b), Output, income and welfare of nations: Concepts and measurement, *Journal of Economic and Social Measurement,* 28, 4, 219-237.

Levin, J. and Nalebuff, B. (1995), An introduction to vote-counting schemes, *Journal of Economic Perspectives*,), 9, 1, 3-26.

McLean, I. and Urken, A. B. (eds.), (1995), Classics of Social Choice, An Arbor, University of Michigan Press.

Miller, N. R. (1987); Voting. In: Eatwell, Milgate and Newman (1987).

Nurmi, H. (1999), *Voting Paradoxes and How to Deal with Them*, Springer Verlag, Berlin.

Robbins, L. (1984), *An Essay on the Nature and Significance of Economic Science*, third ed., New York University Press

Samuelson, P. A. (1972), Pioneers of economic thought, *New York Times*, October 26, 1972.

Sen, A. (1986), Social choice theory, Ch. 22 in: Arrow, K., J. and Intrilligator, M., D. (eds.), *Handbook of Mathematical Economics, vol. III*.

Streeten,P. (1987), Myrdal, Gunnard. In: Eatwell, Milgate, and Newman (1987).

Tabarrok, A. and Spector, L. (1999), Would the Borda count have avoided the Civil War? *Journal of Theoretical Politics*, 11, 2, 261-288.

Tullock G. (2004), Problems of voting, working paper presented at University of Munich.[1]

ABREVIATIONS

AV	approval voting
BC	Borda count
CR	Condorcet rule
CV	cumulative voting
EV	evaluative voting
IRV	instant runoff voting
MR	majority rule
PV	plurality voting
RV	restricted voting
UI	utility index
UV	unrestricted voting
WI	welfare index

[1]